

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 19, 2003, 15:23:30 ; Search time 467 Seconds

(without alignments)
4375.749 Million cell updates/sec

Title: US-09-494-297-2

Perfect score: 3945

Sequence: 1 MKKTRFPNKNTLNTFORVLS.....IAGISLIGMGHTIRIKRD 757

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=xip
-O=/cgn2_1/USPRO.spool/US09494297/unal_13082003.122946.28082/app_query.fasta.1.903
-DB=N.Geneseq.19jun03 -GEMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blusum62 -TRANS=human40.cdi
-LIST=45 -DOCAIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NOR=ext -HEAPSIZE=500 MINLEN=0 -MAXLEN=200000000
-USER=US09494297.ecgn.1.1.304.etunal_13082003.122946.28082 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.19jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3945	100.0	2286	24	ABN69678	Streptococcus poly
2	236.5	6.0	1020	24	ABN69681	Streptococcus poly
3	196	5.0	15614	20	AA112982	Enterococcus faeca
4	196	5.0	15614	20	ABN98777	Enterococcus faeca
5	180	4.6	2127	21	AAA30874	Streptococcus equi
6	178	4.5	7650	20	AAV82019	Moraxella catarrha
7	177	4.5	49617	22	AAF28541	Genomic fragment #
8	176	4.5	2718	20	ABN82021	Moraxella catarrha
9	161	4.1	4161	24	ABQ69442	Listeria innocua D
10	161	4.1	4185	24	ABQ67871	Listeria innocua D
11	161	4.1	4249	20	AAV20201	Listeria innocua D
12	161	4.1	4249	24	ABN98186	E faecalis EP104 g
13	161	4.1	4249	25	ABX61756	Enterococcus faeca
14	161	4.1	4359	20	AAV20200	Enterococcus faeca
15	161	4.1	4359	24	ABN98185	E faecalis EP104 g
16	161	4.1	4359	25	ABX61755	Enterococcus faeca
17	161	4.1	12445	20	AAV13179	Enterococcus faeca
18	161	4.1	12445	24	ABN98974	Enterococcus faeca
19	161	4.1	1163020	24	ABO67197	Listeria innocua C
20	161	4.1	3011208	24	ABO69245	Listeria innocua D
21	159	4.0	4593	24	ABO68111	Listeria monocytog
22	159	4.0	4593	24	ABO70032	Listeria monocytog
23	159	4.0	2944528	24	ABA03041	Listeria monocytog
24	156	4.0	4557	24	ABK75341	Bacillus lichenifo
25	155	3.9	2499	24	ABQ68254	Listeria monocytog
26	155	3.9	2499	24	ABQ68254	Listeria monocytog
27	155	3.9	2944528	24	ABA03041	Listeria monocytog
28	154	3.9	5547	18	AAV68844	Photobacterium lum
29	152.5	3.9	6042	24	ABQ69231	Listeria sp DNA se
30	152	3.9	5547	19	AAV29930	toda gene from the
31	152	3.9	7551	18	AAV68848	Photobacterium lum
32	152	3.9	7551	19	AAV29928	toda gene from the
33	152	3.9	7551	22	AAV35778	Photobacterium lum
34	152	3.9	7551	24	ABF52583	P. luminescens (w-
35	152	3.9	7557	22	AAV58780	Modified Photornab
36	152	3.9	7621	20	AAF58783	Tcda toxin-zeln ER
37	150	3.8	3531	20	AAV77451	S. pyogenes SFEPB-
38	150	3.8	3698	24	AAV35550	Streptococcal fibr
39	149	3.8	2700	21	AAV97541	Human Hsp72 (heat
40	149	3.8	3945	20	AAV7593	S. aureus Sctd DNA
41	149	3.8	4143	25	ABT14918	Pathogen specific
42	149	3.8	6168	20	AAV20228	Enterococcus faeca
43	149	3.8	6168	20	AAV20230	Enterococcus faeca
44	149	3.8	6168	20	AAV20232	Enterococcus faeca
45	149	3.8	6168	24	ABN98213	E faecalis EP123 g

ALIGNMENTS

RESULT 1
ID ABN69678 standard; DNA; 2286 BP.
ID ABN69678;
AC
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 7269.
XX
XX
KW Streptococcus; GAS; GBs; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
XX
PN WO200234771-A2.

XX 02-MAY-2002.
 PD 29-OCT-2001; 2001MO-GE04789.
 XX 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAY-2001; 2001GB-0005640.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Telford H;
 XX MPI; 2002-352536/38.
 DR P-PSDB; ABP29047.
 XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX Claim 7; Page 3879; 4525pp; English.
 PS The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and anti-inflammatory
 CC activity. (I), nucleic acids encoding (I), ABN65044-ABN71556 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX Sequence 2286 BP; 831 A; 347 C; 447 G; 661 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 2,52e-300 Length: 2286
 Score: 3945.00 Matches: 757
 Percent Similarity: 100.008 Conservative: 0
 Best Local Similarity: 100.008 Mismatches: 0
 Query Match: 100.008 Indels: 0
 DB: 24 Gaps: 0
 US-09-494-297-2 (1-757) x ABN69678 (1-2286)
 QY 1 MelLysLysThrArgPheProAsnLysLeuAsnThrLeuAsnThrGlnArgValLeuSer 20
 DB 16 ATATAAAACCAAGGTTCCAAATAGCTTAATCTTAATCTCAAGGATTAAGT 75
 QY 21 LysAsnSerLysArgPheThrValThrLeuValGlyValPheLeuMetIlePheAlaLeu 40
 DB 76 AAAAAGCTCAAAAGCATTTACTGTCTTAACTGAGAGCTCTTTTAAATGATCTTCGCTTG 135
 QY 41 ValThrSerMetValGlyAlaLysThrValPheGlyLeuValGluSerSerThrProAsn 60
 DB 136 GTAACCTTCATGGTGGTCTAGACGCTTTGGTTAGTAGATCTCGAGCCCAAC 195
 QY 61 AlaIleAsnProAspSerSerSerGluTyrArgTrpTyrGlyTyrGluSerTyrValArg 80
 DB 196 GCATTAATTCACATTCACAGTTCGGAATACAGATGATGATGATATGATCTTATGTAGA 255
 QY 81 GlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGluGly 100
 DB 256 GGGCATCCATATTATTAACAGTTTAGAGTAGACACAGCATTTAAAGGTTAACTTAGAAGA 315

QY 101 SerArgSerTyrGlnValTyrCysPheAsnLeuLysLysAlaPheProLeuGlySerAsp 120
 DB 316 AGTAGAGGTATACAGAGTTTATGCTTTAATTTAAAGAAACAGATTCTCTCGGATCAGAT 375
 QY 121 SerSerValLysLysTyrTyrLysLysHisAspGlyIleSerThrLysPheGluAspTyr 140
 DB 376 AGATGTGTAAAGAGGTGTATTAAGAAAAACATGATGGAATCTCTCAAAAAATTTGAAGATTAT 435
 QY 141 AlaMetSerProArgIleThrGlyAspGluLeuAsnGlnLysLeuValGlyAlaValMetTyr 160
 DB 436 GCGATGAGCCCTAGCAATTTACGGAGATGAGCTTAATACAGAGTTAGAGCTGTATATGTAT 495
 QY 161 AsnGlyHisProGlnAsnAlaAsnGlyIleMetGlnGlyLeuGluProLeuAsnAlaIle 180
 DB 496 AATGACATCCACAAATATGCAATATGATATATGGAAGCTTTGAACCTTGAATGCTATC 555
 QY 181 ArgValThrGlnGluAlaValTyrTyrTyrSerAspAsnAlaProIleSerAsnProAsp 200
 DB 556 AGAGTTACACAGAGGCGGTATGTACTATTCGATATGCTCTATTTCTTAATCCAGAT 615
 QY 201 GluSerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuMet 220
 DB 616 GAAAGTTTAAAGGAGTCAGAAAGTAACTTGTTAGTCTCTCAATTATCTTTGATG 675
 QY 221 ArgGlnAlaLeuLysGlnLeuLysProAsnLeuAlaThrLysMetProLysGlnVal 240
 DB 676 CGTCAACCTTTGAAGCAACAGATGATCGCAATTTGGCACTAAATGCCAATAAACAAGT 735
 QY 241 ProAspAspPheGlnLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsnLys 260
 DB 736 CCGGATATTTTACGCAAGTATTTTGAAGTCTGAGCAACAGGAGATTAATTAATAA 795
 QY 261 GlyTyrGlnAsnLeuLeuSerGlyGlyLeuValProThrLysProProThrProGlyLysP 280
 DB 796 GGATACCAAAATCTTTTGAAGTGTGTTAGTTTCTCTTAACCAACCACTCCAGGAGAC 855
 QY 281 ProPheMetProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIle 300
 DB 856 CCACCAATGGCTCCAAATCAACCTCAACGCTCACTACTTTTAAAGATGTGCTATA 915
 QY 301 GlyAspTyrSerLysLeuLeuGluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsn 320
 DB 916 GGTGATTACTCTAAATGCTTGAAGGTGCAACATTAAGCTTGACAGGGGATTAACGTGAAT 975
 QY 321 SerPheGlnAlaArgValPheSerSerAsnAspIleGlyLysArgIleGluLeuSerAsp 340
 DB 976 AGTTTTCACCGAGAGTGTGTAGCAGTAATGATATTTGGAGAAAGATTTGAATCATCAGAT 1035
 QY 341 GlyThrTyrThrLeuThrGluLeuAsnSerProAlaGlyTyrSerIleAlaGluProIle 360
 DB 1036 GGAACCTATACCTTAACCTGAATGAATTCACAGCTGTATATGATATCGCAGAGCCATC 1095
 QY 361 ThrPheLysValGluAlaGlyLysValTyrThrIleLeuAspGlyLysGlnIleGluAsn 380
 DB 1096 ACTTTTAAGGTTGAAGCTGCAAGCTTAACATTAATTAATGAGAAACAGATTGAAAT 1155
 QY 381 ProAsnLysGlnIleValGluProTyrSerValGluAlaTyrAsnAspPheGluGluPhe 400
 DB 1156 CCAATATAAGATAGTAGAGCCTTACTCGATGCAACCATATATATATTTTGAAGAAATTT 1215
 QY 401 SerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAsnLysAsnGlySer 420
 DB 1216 AGCGTTTAACTACACAAACCTATGCAAAATTTTATATATCAAAAAATTAATTAAGTAGT 1275
 QY 421 SerGlnValValTyrCysPheAsnAlaAspLeuLysSerProProAspSerGluAspGly 440
 DB 1276 TCACAGGTGTCTATGTCTTAAATGAGATCTAAATCTCCACAGAGCTCTGGAAGAGGT 1335
 QY 441 GlyLysThrMetThrProAspPheThrThrGlyGluValLysTyrThrHisIleAlaGly 460
 DB 1336 GGGAAACAAATGACTCCAGACTTTTACACAGGAGAACTAAATATACACTCATATTGCAAGT 1395


```

QY 479 LeuLysHisIleLysLysValIleGluLysGlyTyrArgGluLysGlyGlnAlaIleGlu 498
Db 262 -----AAAGTGATCAAAATGCAAAATGCAATTTGAT 297
QY 499 TyrSerGlyLeuThrGluThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrTyrPhe 518
Db 298 TTTTCAGAACTTCT-----TTGCAAAACCCAGCTGTTTATTATTATAC 339
QY 519 ThrAspSerIleGlu-----LeuAspLysAspLysLeuLysAspTyrHisGlyPheGlyAsp 537
Db 340 AAAGTAACTGAGAGAGAGATAGATAAAGTTCTGCTGCTTCTTAT----- 384
QY 538 MetAsnAspSerThrLeuAlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsn 557
Db 385 -----GATACAACTCTTACACTGTTCAAGTTTCATCTCTTGAGATGAAGACCAACA 438
QY 558 ProProGlnLeuThrAspLeuAspPhePheIleProAsnAsnLysTyrGlnSerLeu 577
Db 439 AAACCACTAGCTTACT-----TATATT 459
QY 578 IleGlyThrGlnThrHisProGluAspLeuValAspIleIleArgMetGluAspLysLys 597
Db 460 GTTGCTTATATA-----GAAGTAGTAG 483
QY 598 GluValIleProVal-----ThrHisAsnLeuThrLeuArgLys 610
Db 484 -----GTGCCAATTCAGTTCAAAATAGCTTAGATTCTATCATTAACGGTGAAGAAA 537
QY 611 ThrValThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGluLeuLys 630
Db 538 AAAGTTTCAGTACCGGTGGAGATCGCTTAAAGATTTTATTTTGTCTGACGTTTAAAA 597
QY 631 AsnAspLysGlnLeuLeuLeuSerGlnThrValLysThrAspLysThrAsnLeuGluPhe 650
Db 598 GCAAATCAGATATATTAAGGCGTCAAGAAAAGTATGATTAAGAGCAACT----- 648
QY 651 LysAspGlyLysAlaThrIle-----Asn 658
Db 649 AAAGTGTCACACTCTCTTCAACAGAGCTAGTATAGATCAACTCTATCATTTTACC 708
QY 659 LeuLysHisGlyLysLeuThrLeuThrGlnGlyLeuProGluLysTyrSerTyrLeuVal 678
Db 709 TTGAAGATGATTAATCAATCAAAAGTCAAAATCTTCACTAGCTGATGATTTGTTGTC 768
QY 679 LysGluThrAsp-----SerGluGlyTyr-----LysValLysValAsnSerGlnGlu 694
Db 769 ACTGAAGACGATTAACAATCAAGAAATATACAAACCAAGCTGGAAGTTAGTCTCTCAAGAT 828
QY 695 ValAla-----AsnAlaThrValSerLysThrGlyIleThrSerAsp 708
Db 829 GGAGCTGTAAAAAATATCGCAGCTAATTCMACTGACACAGACAGCACTACTGATAAAGAT 888
QY 709 GluThrLeuAlaPheGluAsnAsnLysGluProValValProThrGlyValAspGlnLys 728
Db 889 ATGACCATTAATCTTCAAAATTAATAAAGACTTTGAAGTCCAAACAGAGTAGCAATGACT 948
QY 729 IleAsnGlyTyrLeuAlaLeuIleValIleAla 739
Db 949 GTGGCACATATATTTCTTAGCAATTTAGCA 981

```

```

XX OS Enterococcus faecalis.
XX PN WO9850555-A2.
XX PD 12-NOV-1998.
XX PF 04-MAY-1998; 98MO-US08985.
XX PR 14-NOV-1997; 97US-0066009.
XX PR 06-MAY-1997; 97US-0044031.
XX PR 16-MAY-1997; 97US-0046655.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Barash SC, Dillon PJ, Kunsch CA;
XX WPI; 1999-045171/04.
XX PT New isolated Enterococcus faecalis polynucleotides and polypeptides
XX PT - used to develop products for the detection of Enterococcus and for
XX PT use in vaccines for prevention or attenuation of Enterococcus
XX PT infection.
XX PS Claim 1; Page 419-427; 2084pp; English.
XX CC A computer readable medium has been developed which has recorded on it
XX CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
XX CC AAX12938 to AAX13919 represent these nucleotide sequences which are
XX CC primary nucleotide sequences, also known as contigs. The computer based
XX CC system can identify fragments of the Enterococcus faecalis genome with
XX CC commercial importance. The products can be used to detect the presence
XX CC of Enterococcus faecalis in samples. They can also be used for
XX CC diagnosing Enterococcal infection in an animal and monitoring
XX CC progression of disease, and for identifying agents which can be used to
XX CC modulate the growth or pathogenicity of Enterococcus faecalis, or
XX CC another related organism, in vivo or in vitro. In particular the
XX CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
XX CC can be used in vaccines to prevent or attenuate an Enterococcal
XX CC infection.
XX SQ Sequence 15614 BP; 5535 A; 2567 C; 3169 G; 4326 T; 17 other:

Alignment Scores:
Pred. No.: 8.13e-05 Length: 15614
Score: 196.00 Matches: 165
Percent Similarity: 35.57% Conservative: 89
Best Local Similarity: 23.11% Mismatches: 245
Query Match: 4.97% Indels: 216
DB: 20 Gaps: 41

US-09-494-297-2 (1-757) x AAX12982 (1-15614)
QY 88 PheArgValAlaHisAspLeuArgValAsnLeuGlySer-----ArgSerTyrGln 105
Db 9047 TTGAGCTTCAAAACATTA-CCAGAAGTGTCTCTCAACGCCCTCAATTAATATGATGAG 9105
QY 106 ValTyrCysPheAsnLeuLysAlaPheProLeuGlySerAspSerSerValLys 125
Db 9106 ATTTAT-----TTGGGTAAAGGCAAGAAATTCAAAT 9138
QY 126 TrpTyrLysHisAspGlyIleSerThrLysPheGluAsp----- 139
Db 9139 CATTATCAAGTACT-----ATTCAACAGAGCTCAAGAAACTGATTTTGG 9192
QY 140 TyrAlaMetSerProArgIleThrGlyAspGluLeuAsnGlnLysLeuArgAlaValMet 159
Db 9193 TATCAATGATGATGGAGACAGTTTCAGCCATTCACAGGCCCTGAAAAAGTTGAT 9252
QY 160 TyrAsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluProLeuAsnAla 179
Db 9253 TTT-----GGGTTCTTCCGGGAAAGCACCT-----GGC 9282

```

RESULT 3
 AAX12982
 ID AAX12982 standard; DNA; 15614 BP.
 XX
 AC AAX12982;
 XX
 DT 19-MAR-1999 (first entry)
 DE Enterococcus faecalis genome contig SEQ ID NO:45.
 XX
 XX Enterococcus faecalis; contig; detection; Enterococcal infection;
 KW vaccine; attenuation; computer readable medium; ds.

OY 180 IleArgValThrGlnGluAlaValTyrTyrSerAspAsnAlaProIleSerAspPro 199
 Db 9283 GTGAGTTAAACGTAAGAAAAATCTGGAGAGATGATCAAGACCCGACAAAGTCGGCCA 9342
 OY 200 AspIuSerPheLysArgIuSerGluSerAsnLeuValSerThrSer----- 215
 Db 9343 GATAAATGTGATTTGAAATATAGTAAAGCAAGTAAGTACACAGCAACTGGCAAACT 9402
 OY 216 -----GlnLeuSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeu 232
 Db 9403 GGGTATATTAATATATCAAAACAGAAATATGATACCAAGTAAGTGGAGCGCAAAAT 9462
 OY 233 AlaThrIysMetProLysGlnValProAspSphPheGln-----LeuSerIlePheGlu 250
 Db 9463 GTAAACCAACTTTCACAAACCGCGATGAAGCTATCAAGAGTTCTTGGCTTCCCAA 9522
 OY 251 SerGluAspLysGluAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGlyLeu 270
 Db 9523 TACAAACATCAAGCAAGCAAGCTTTCAT-----TATCAACACACCCGGAATTAGCA--- 9573
 OY 271 ValProThrIysProProThrProGlyAspProPheMetProProAsnGlnProGlnThr 290
 Db 9574 GTTCTCGTTACAGTCAAGAAAAATCGACATACCTAGTGAAGAAACACCAACAGCTTC 9633
 OY 291 ThrSerVal-----LeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuLeu 307
 Db 9634 AAGCATTAAGTAAAGTAAATCAAAATCTTCTCCAGGTGAG-----AAAACTTA 9687
 OY 308 GlnGluValaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgValPhe 327
 Db 9688 GTGGAGCCCGCTTGTGAATTCAGTGAATAAATGTT-----CAACAAACATTAATG 9738
 OY 328 SerSerAsnAsp-----IleGlyGlnArgIleGluLeuSerAspGly---Thr 342
 Db 9739 GACATTAAGATGTGTACTTATCTCTCCCAAGATGCGCTTACAAAAGGGGAGACGC 9798
 OY 343 TyrThrLeuThrGluLeuAsnSerProAlaGlyTyrSerIleAlaGluProIleThrPhe 362
 Db 9799 TATCATTAAGTAAAGTAAAGACACCTGACGACATGATGACAAAGAAACGACTTGG 9858
 OY 363 LysValGlnAla-----GlyLysValTyrThrIleLeuAspGlyLysGlnIleGlu 379
 Db 9859 CAATTTAGGTGACTGAGCAAGCAAGTA-----AGCATCGATGACAGAGTGCAC 9912
 OY 380 AsnProAsnLysGluIleValGluProTyrSerValGluAlaTyrAsnAspPheGluGlu 399
 Db 9913 ACCCAATCAAGATTATT-----CCATTGGAAATGAA-----AATTAATTTCTTCT 9960
 OY 400 PheSerVal-----LeuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLys 415
 Db 9961 TTGCCAATCAAGATTAGAAAAATACACCATGCATAAT-----GGCATA 10002
 OY 416 AsnLysAsnGlySerSer-GlnValValTyrCysPheAsnAlaAspLeuLysSerProPr 435
 Db 10003 CAAGTGAACCTTAGCAGAGCGCACTTTTCCGTTCGAAGAA----- 10042
 OY 435 AspSerGluAspGlyLysThrMetThrProAspPheThrThrGlyGluValLysTyr 455
 Db 10043 -----AAAATGCTGACAGAAAGTACCAAACTGCGCAACTGCAAAA 10083
 OY 455 rThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProArgAspThrAspPr 475
 Db 10084 AACGATACTACAGAGATGAGCTATTT----- 10111
 OY 475 AspThrPheLeuLysHisIleLysValIleGluLysGly---TyrArgGluLysGlu 494
 Db 10112 -----AAATATGTAACCTGTGATGATGCA----- 10138
 OY 494 yGlnAlaIleGluTyrSerGlyLeuThrGluThrGlnLeuArgAlaAlaThrGlnLeuAl 514
 Db 10139 -----ATGTGCAACATCAGCA-----CCATTAGCTACGACACTCTTGTGG 10182
 OY 514 aIleTyrTyrPheThrPheSerAlaGluLeuAspLysAspLysLeuLysAspTyrHisGlu 534

Db 10183 AAATTAATCAATTACT-----GTTGATTA----- 10207
 OY 534 yPheGlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGluTyrAlaGlu 554
 Db 10208 -TATGGGAAA-----ATTCACTATGCAAG 10230
 OY 554 n-----AspSerAsnProProGlnLeuThrAspLeuAspPhePheIleProAsnAs 571
 Db 10231 CAAAATATTTAGCAAAATGCGCAGATGAGCA----- 10264
 OY 571 nasLysTyrGlnSerLeuIleGlyThrGlnThrPisProGluAspLeuValAspIleI 591
 Db 10264 ----- 10264
 OY 591 eArgMetGluAspLysGluValIleProValThrHisAsnLeuThrLeuArgLys-- 610
 Db 10265 -----CTGACACATCAAAATAATTGAAACCTTT 10293
 OY 611 -----ThrValThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGlu 628
 Db 10294 TGACTTAACAGTTATATAAAAAAGCGATATATCAGACG-----CC 10332
 OY 628 uLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThrAspLysThrAsnLe 648
 Db 10333 ACTTAAGAGACGAAATTCGTTTACAGCA-----CCAGATACGATAT 10377
 OY 648 uGluPhe---LysAspGlyLysAlaThr-----IleAsnLeuLysHisGlu 662
 Db 10378 TGAATTACCAAAAGTGCAGAAAGACGATACTTGTGTTTGAATAAAGTAAACGAG 10437
 OY 662 yGlu---SerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGluTh 681
 Db 10438 GAAATATGCTTACACAGAAACCTTACCGCAGAAAGATATCAGGGCTTAAAGAACCAAT 10497
 OY 681 rAsp-----SerGlnGlyTyrLysValLysValAsnSerGlnGluAlaAlaAsnAl 698
 Db 10498 CGAATTATTAATCTCTGAAAGATGTTACGATACGATACATGCGCAAAAGTACGAGAT 10557
 OY 698 aThrValSer-----LysThrGlyIleThrSerAspGluThrLeuAlaPheGluAs 715
 Db 10558 TTTAATTTCTGAGAGAAAGAAATATCAAAATTAATCTTAACGTTACG-----AACCA 10608
 OY 715 nasLysGluProValValProThrGlyValAspGlnLysIleAsnGlyTyrLeuAlaLe 735
 Db 10609 AGCAATAGGTTCTCTTACCTGAAACGTCGGCATAGAGACGTTGTGCTTACTGATAGC 10668
 OY 735 uIle-----ValIleAlaGlyIleSerLeuGlyIle 745
 Db 10669 GATTATGATCAATTCGTGATAGCGGCTTATCTCTTATT 10708
 Db 10708

RESULT 4
 ABS98777 standard; DNA: 15614 BP.
 ABS98777;
 18-DEC-2002 (first entry)
 Enterococcus faecalis contig sequence #45.
 Computer readable medium; Enterococcus faecalis; microbe; growth;
 pathogenicity; vaccine; resistance; Enterococcal infection; commercial;
 therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;
 biochip technology; antibacterial; modulator of nucleic acid expression;
 contig; ds.
 Enterococcus faecalis.
 US2002120116-A1.
 29-AUG-2002.

PF	04-MAY-1998:	98US-0070927.	
XX			
PR	04-MAY-1998:	98US-0070927.	
XX			
PA	(KUNSCH C. A.		
PA	(DILLON P. J.		
PA	(BARASH S.		
XX			
PI	Kunsch CA, Dillon PJ, Barash S;		
XX			
DR	WPI: 2002-750065/81.		
XX			
PT	Computer readable medium having recorded on it a <i>Enterococcus faecalis</i>		
PT	nucleotide sequence useful for detecting diseases related to		
PT	<i>Enterococcus</i> infections in animals		
XX			
PS	Claim 1; Page -; 119pp; English.		
XX			
CC	The present invention relates to a new computer readable medium with an		
CC	<i>Enterococcus faecalis</i> nucleotide sequence. The invention is useful to		
CC	diagnose the presence of <i>E. faecalis</i> in a sample or determining the		
CC	presence of a specific microbe in a sample. The invention is also useful		
CC	for modulating the growth or pathogenicity of <i>E. faecalis</i> in a vaccine		
CC	to confer resistance to <i>Enterococcal</i> infection, for commercial,		
CC	therapeutic and industrial purposes, and for fermenting a particular		
CC	sugar source or to produce a particular metabolite. The invention is		
CC	useful for detecting diseases related to <i>Enterococcus</i> infections in		
CC	animals, and for detecting <i>E. faecalis</i> using biotech technology. The		
CC	present nucleic acid sequence represents an <i>Enterococcus faecalis</i> contig		
CC	DNA sequence of the invention.		
CC	Note: The sequence data for this patent did not form part of the		
CC	printed specification but was obtained in electronic format directly		
CC	from USPTO at http://seqdata.uspto.gov .		
XX			
SQ	Sequence 15614 BP; 5535 A; 2567 C; 3169 G; 4326 T; 17 other:		
	Alignment Scores:		
	Pred. No.:	8.13e-05	Length: 15614
	Score:	196.00	Matches: 165
	Percent Similarity:	35.57%	Conservative: 89
	Best Local Similarity:	23.11%	Mismatches: 245
	Query Match:	4.97%	Indels: 216
	DB:	24	Gaps: 41
US-09-494-297-2 (1-757) x ABS98777 (1-15614)			
OY	88 PheArGyAlAlAlAlHsAspLeuArGyAlAsnLeuGlnGlySer-----ArgSerTyrGln	105	
DB	9047 TTCAGGTTCACAAACATTAA-CCAGAGAGTCGCGTACACAGCGCTACAAATTAATGTAAGAG	9105	
OY	106 ValArgCysPheAsnLeuLysLysAlaPheProLeuGlnGlySerAspSerValLysLys	125	
	:::		
DB	9106 ATTATAT-----TTGGGTAAAGGCCAAGAAATTCAAATT	9138	
OY	126 TrpTyrLysLysHsAspGlyLleSerThrLysPheGluAsp-----	139	
DB	9139 CATATATCAAGTACG-----ATTCAAAACAGAGTCAGAAACTCAAACTGATTTTGG	9192	
OY	140 TyrAlaMetSerProArgLleThrGlyAspLeuLeuAsnGlnLysLysLeuAlaGlyAlaMet	159	
DB	9193 TATCAAAATGAATGTCGGACCAACGTTTCAGCCATTAGCCACAGGCCCTCGAAAAAGTTGAT	9252	
OY	160 TyrAsnGlnGlyHisProGlnAsnAlaAsnGlnLysLysLeuGlnLysLeuAlaAsnAla	179	
DB	9253 TTT-----GGGGTTCCTTCGGGAAAGCACT-----GGC	9282	
OY	180 IleArgValThrGlnGluAlaValTrpTyrTyrSerAspAlaProIleSerAspPro	199	
DB	9283 GTGAAGTTAAACGTGAAATAAAATCTGGGAGACAGATGATCAAAAGCCGACAAAGTCGGCA	9342	
OY	200 AspGluSerPheLysArgGlnGlySerGluSerLysLeuValSerThrSer-----	215	
DB	9343 GATATATGATTTATGAATTAATGTAAGAAAGCAAGCACTGACACAGGCCAACTGCGCAACT	9402	

```

OY 554 n-----AspSerAsnProGlnLeuThraspleuAspPheIleProAsnAs 571
DB 10231 CAAAAATATGAGAAATGCGCCAGATGAGACA----- 10264
OY 571 nasnlystYrGlnSerLeuIleGlyThrGlnTrpHisProGluAspLeuValAspIleI 591
DB 10264 ----- 10264
OY 591 eaRgMeCluAspLysLysGluValIleProValThrHisAsnLeuThrLeuArgLys-- 610
DB 10265 -----CTGACACATCAAAATAATTAATTTGAACACCTTT 10293
OY 611 -----ThrValThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleG 628
DB 10294 TGACTTAACAGTATTAATAAAAGCCGATATCAGACG-----CC 10332
OY 628 uLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThrAspLysThrAsnle 648
DB 10333 ACTTAAGAGCAGCAAAATCCGTTAAACAGCA-----CCAGATACGATAT 10377
OY 648 uGluPhe---LysAspGlyLysAlaThr-----IleAsnLeuLysHisG 662
DB 10378 TGAATTAACCAAAAGATGCAAAAGAACGATCTTTGTTTGAACCTTAACCAACAG 10437
OY 662 YGlu---SerLeuThrLeuGlnGlyLeuProGlnGlyLysSerYrLeuValLysGluTh 681
DB 10438 GAAATATGTTTAAACAGAACCTTTACGCCAGAAAGATATACAGGGGTTTAAAGAACCAAT 10497
OY 681 rAsp-----SerGlnGlyTyrLysValLysValAsnSerGlnGluValAlaAsnAl 698
DB 10498 CGAATTATATATGTCGAAGATNGTTCAGTCACGATAGATGGGAAAGATAGCAGATGT 10557
OY 698 aThrValSer-----LysThrGlyIleThrSerAspGluThrLeuAlaPheGluAs 715
DB 10558 TTTAATTTCTGAGAGAGAAATCAAAATTAATTAATTAACCTTACACCTTACG-----AACCA 10608
OY 715 nasnLysGluProValValProThrGlyValAspGlnLysIleAsnGlyTyrLeuAlaLe 735
DB 10609 AGCAAAAGTCTCTTACCTGAACACTGGTGCGATAGACGCTTGCTTTTACTTGATAGC 10668
OY 735 uIle-----ValIleAlaGlyIleSerLeuGlyIle 745
DB 10669 GATTAGTACATTCGTCGATAGCGGCTTATCTCTTTATT 10708

```

```

RESULT 5
AAA30874
ID AAA30874 standard; DNA: 2127 BP.
XX AAA30874;
XX
XX 19-SEP-2000 (first entry)
DE Streptococcus equi fibronectin binding protein, FNZ, coding sequence.
XX
XX Fibronectin binding protein: SFS; vaccine; horse; strangles; therapy;
KW equine upper respiratory tract disease; S. equi infection; FNZ; ds.
XX
OS Streptococcus equi.
XX
XX Key location/Qualifiers
XX CDS 108..1901
XX FT /*tag=a
XX FT /product= FNZ
XX
XX MO2000037496-A1.
XX
XX 29-JUN-2000.
XX
XX 21-DEC-1999; 99WO-SE02448.
XX
XX 22-DEC-1998; 98SE-0004491.
XX
XX (GUS5/) GUS5 B.

```

```

PA (LIND/) LINDMARK H.
PA (JACO/) JACOBSSON K.
PA (FRK/) FRKBERG L.
XX
XX Guss B, Lindmark H, Jacobsson K, Frykberg L;
DR WPI: 2000-442641/38.
DR P-PSDB: AAV90257.
XX
XX New protein useful for preparation of vaccines for treatment of
PT strangles caused by Streptococcus equi infection, is able to bind to
PT mammalian fibronectin -
XX
PS Disclosure: Page 17a-17b; 34pp; English.
XX
XX This sequence encodes the Streptococcus equi fibronectin binding
CC protein FNZ. The FNZ sequence was used to isolate the S. equi fibronectin
CC binding protein of the invention, designated SFS. SFS binds specifically
CC to mammalian fibronectin or its analogues or fragments. The protein, its
CC analogues or fragments may be used for the preparation of a vaccine that
CC protects horses against strangles (a world-wide distributed and serious
CC disease of the equine upper respiratory tract) caused by S. equi
CC infection. The antibody and/or antiserum may also be used for the
CC prophylactic or therapeutic treatment of S. equi infection in mammal,
CC especially horses. The use of vaccines containing the fibronectin binding
CC protein provides a more effective protection against S. equi infections,
CC with fewer side effects.
XX
SQ Sequence 2127 BP; 718 A; 415 C; 466 G; 528 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.000124 Length: 2127
XX Score: 180.00 Matches: 150
XX Percent Similarity: 34.64% Conservative: 98
XX Best Local Similarity: 20.95% Mismatches: 270
XX Query Match: 4.56% Indels: 198
XX DB: 21 Gaps: 33
XX
US-09-494-297-2 (1-757) x AAA30874 (1-2127)
OY 68 SerGluTyrArgTrpTyrGlyTyrGlnSerTyrValArg---GlyHisProTyrTyrLys 86
DB 201 GCAGAGCAGCCTTATATATGCGGAAATGATGAGACAGCAAAAGTCCCATATTT--- 257
OY 87 GlnPheArgValAlaHisAspLeuArgValAsnLeuGlnGlySerArgSerTyrGlnVal 106
DB 258 ---TTGTACGATATCCCTAAATAATGCTCCAAAGCCTGAATTTAAAGACGAGATGTGTT 314
OY 107 TyrGlyPheAsnLeuLysLysAlaPheProLeuGlnGlySerAspSerValLysLysTrp 126
DB 315 TATTGCTTTAAACAAATAATGTATGTGGCAGATCAATGGGAATCTATATACAGCAATTTT 374
OY 127 -----TyrLysLysHisAspLysIleSer 134
DB 375 AATGCATCAGATTCATATACGATTTACCTGATATGAGAAAAAACTAGAGATATGAT 434
OY 135 ThrLysPheGluAspTyrAlaMetSerProArgIleThrGlyAspGluLeuAsnGlnLys 154
DB 435 GGTATATTTAAACAATATGCTCCAGATTAACAATAAAGATATATGATATTCAGAGTCT 494
OY 155 LeuArgAlaValaMetTyrAsnGlnLysHisProGlnAsnAlaAsnGlyIleMetGluGly--- 173
DB 495 TTGGTGCAAGTTTAAATGATGATATGACCCCACTAAACAGTCAACAATCAACTAGCTAC 554
OY 174 ---LeuGlnProLeuAsnAlaIleArgValThrGlnGluAlaValTrpTyrTyrSerAsp 192
DB 555 CATTTAATATATGATTTCTTGAAGAAAGTATACATATTCAGCATTTGGTATTTTATGAT 614
OY 193 AsnAlaProIleSerAsnProAspGlnSerPheLysArgGluSerGlnSerAsnLeuVal 212
DB 615 -----AGTTTAAACAAGAAATACCTTAAGATACGGGCTTATTAACCTTA--- 659
OY 213 SerThrSerGlnLeuSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeu 232

```



```

Db 660 -----AACGATATGAGAAAAAGCTTTAGATTATTTAAACAGTAAAGAGAG 707
QY 233 AlaThrLysMetProLysGlnValProAspAspPheGlnLeuSerIlePheGlnSerGlu 252
Db 708 GATCTTAAGCTT---AAATCAGAGCAGAGTAAATACCATGGATATTATTTATGTTTATCAA 764
QY 253 AspLysGlyAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGlyLeuValPro 272
Db 765 AGTGGGGGGCATGACCATATGAAAGATTACCAAAATCTTCTGGCTTACTTAAATTCCT 824
QY 273 ThrLysProProThrProGlyAspProPheProMetProProAsnGlnProGlnThrSer 292
Db 825 AAAGAACCCTGTAAGCTT----- 842
QY 293 ValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuLeuGlnGlyAlaThrLeu 312
Db 843 -----CAGCTAGGCTGTTTAACT----- 860
QY 313 GlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIle 332
Db 861 -----GGACATATATGAAATGATTAAGCGCCTTGAAGAGAGATCATCAGTTCA 911
QY 333 GlyGlnArgIleGlnLeuSerAspGlyThrTyrThrLeuThrGlnLeuAsnSerProAla 352
Db 912 CAAGAAACTAATGAA-----GATGTTAAGAAAGAACTT-----ATAGCTTTCATGGA 959
QY 353 GlyTyrSerIleLeuGln-----ProIleThr-----PheLysValGlnAlaGly 367
Db 960 GGCATCTCAGAAAGCCAGGCAACGAGATCCTTTGCGAGAGATTGAAGGGTGAAGCTGGGT 1019
QY 368 LysValTyrThrIleLeuAspGlyLysGlnIleGlnAsnProAsnLysGlnIleValGln 387
Db 1020 GCACATGATACA-----CTCAAAAGCCTAATGATTCATCTG----- 1055
QY 388 ProTyrSerValGlnAlaTyrAsnAspPheGlnGluPheSerValLeuThrThrGlnAsn 407
Db 1056 ---CAAGCTTTAAGCGGTAAAC-----TCTCCTAATAGTAAACAAAC 1097
QY 408 TyrAlaLysPheTyrTyrAlaLysAsnLysAsnGlySerSerGlnValValTyrCysPhe 427
Db 1098 TATGTAGT-----ACCGAAGCATATCATGTGATCATCAGCATCTTGAG----- 1142
QY 428 AsnAlaAspLeuLysSerProAspSerGluAspGlyLysThrMetThrProAsp 447
Db 1143 GAAGACGAGATACCTAACCCACTGATCATCTGATGCGGCGGCTAGCTGAG--- 1199
QY 448 PheThrThrGlyGlnValLysTyrThrHisIleAlaGly-----Arg 461
Db 1200 ---ACCGATGAAATACTATAAACCCCTCATCTGATGCGGCGGCTAGCTGAG 1256
QY 462 AspLeuPheLysTyrThrValLysProArg-----Asp 472
Db 1257 GAATTCAGAGAAAGCAGCACTTAACAGCAACCAACCGCGGCGGAGCAGCATCAGCAG 1316
QY 473 ThrAspProAspThrPheLeuLysHisIleLysValIleGlnLysGlyTyrArgGln 492
Db 1317 ACACAGAGAGATACA-----CAA 1334
QY 493 LysGlyGlnAlaIleGlnTyrSerGlyLeuThrGlnThrGlnLeuArgAlaIleThrGln 512
Db 1335 AAAGGATGCTGTCGACATCTGCTGCTACATCATGAGTCAAGAAACACCAAAAGCCGAG 1394
QY 513 LeuAlaIleLysTyrThrThrAspSerAlaGlnLeuAspLysAspLysAspTyr 532
Db 1395 GTATGATTTGGTGTAGAGCAACACATCAGACAAACAGACACACACAAAA----- 1448
QY 533 HisGlyPheGlyAspPheLysAsnSerThrLeuAlaValAlaLysIleLeuValGlnTyr 552
Db 1449 ---GGCATGCTCTGACATCTGCGGCTACTATGAG----- 1481
QY 553 AlaGlnAspSerAsnProProGlnLeuThrAspLeuAspPhePheIleProAsnAsnAsn 572

```

```

Db 1482 TCAGAGCAGCACTAAGAAACCTGAGCTC----- 1508
QY 573 LysTyrClnSerLeuIleGlyThrGlnThrPheIleProGlnAspLeuValAspIleIleArg 592
Db 1509 -----ATGATGCTGGCTAG----- 1523
QY 593 MetGlnAspLysLysGlnValIleProValThrHisAsnLeuThrLeuArgLysThrVal 612
Db 1524 -----GGACAAATCATCGACTCTCTGAAAC-----ACCCAAATCAGATG 1565
QY 613 ThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGlnIleGlnLeuLysAsnAsn 632
Db 1566 TCTGGCGAGCTGCTGTCAGCTAGC-----GTAATTGAG----- 1598
QY 633 LysGlnGlnLeuLeuSerGlnThrValLysThrAspLysThrAsnLeuGlnPheLysAsp 652
Db 1599 -----GATACCAAGACAGCTGAGATTAATCATTTGGTGG 1631
QY 653 GlyLysAlaThrIleAsnLeuLysHisGlyGlnSerLeuThrLeuGlnGlnLeuProGln 672
Db 1632 CAAGACAAATCATCGACTCTCTGAGAGATACAGCCGGGTATGTTGTT---CAATCT 1688
QY 673 GlyTyrSerTyrLeuValLysGlnThrAspSerGlnGlyTyrLysValLysValAsnSer 692
Db 1689 GAGGACCTCAATTTGCGAAGACCCAGAACCCGACCTAAGCCTTAACCTGACCT 1748
QY 693 GlnGlnValAlaAsnAlaThrValSerLysThrGlyIleThrSerAspGlnThrLeuAla 712
Db 1749 GCGCCAAATTTTAAT-----GACGAA----- 1769
QY 713 PheGlnAsnAsnLysGlnProValIleProThrGlyValAspGlnLys----- 1788
Db 1770 ---AAACCTAACAAGGACCTCATCTCCACAGACAAAGTATGAAGCAATCACCCCTA 1826
QY 729 ---IleAsnGlyTyrLeuAlaLeuIleValIleAlaGlyIleSerLeu 743
Db 1827 AGCATCATCGTGCATCAATGCAATGCTGCTGCTATGTTGTTCTTA 1874

RESULT 6
AAV82019
ID AAV82019 standard; DNA: 7650 BP.
AC AAV82019;
XX 21-JUN-1999 (first entry)
DE Moraxella catarrhalis 4223 lfr region.
XX
KW Lactoferrin receptor; lactoferrin binding protein; LBP1; LBP2;
KW lbpA gene; lbpB gene; ORF3; infection; otitis media; sinusitis;
KW conjunctivitis; pneumonia; bronchitis; tracheitis; emphysema;
KW diagnosis; therapy; vaccine; Branhamella catarrhalis; ss.
XX
OS Moraxella catarrhalis.
XX
FH key location/Qualifiers
FT -35_signal 64..69 /*tag= a
FT -10_signal 98..103 /*tag= b
FT RBS 127..136 /*tag= c
FT CDS 141..2837 /*tag= d
FT /gene= lbpB
FT /product= LBP2
FT /transl_except= (pos:2031..2032, aa:Asn)
FT /note= "this codon has an apparent 1 codon deletion
FT which alters the reading frame"
FT /transl_except= (pos:2036..2039, aa:Ala)
FT /note= "this codon has an apparent 1 codon
FT insertion which alters the reading frame"
FT /note= "lbpB is specifically claimed in Claim 8;

```


FT misc_feature 632 encodes AAM89413"

FT /tag= e

FT /note= "this base is not present in the sequence

FT given in the specification, the nucleotide

FT has been included to maintain the reading

FT frame as part of a codon (AAR) for Lys"

FT -35_signal 2973..2978

FT /tag= f

FT 2990..2995

FT /tag= g

FT 3006..3015

FT /tag= h

FT 3021..6023

FT /tag= i

FT /gene= lbpA

FT /product= lbp1

FT /note= "specifically claimed in Claim 8; encodes

FT AAM89414"

FT CDS 3066..6023

FT /tag= j

FT /gene= lbpA

FT /product= lbp1

FT /note= "specifically claimed in claim 8; encodes

FT AAM89415"

FT 6025..7650

FT /tag= k

FT /gene= orf3

FT /note= "specifically claimed in Claim 7; encodes

FT AAM89416"

FT W09855606-A2.

PD 10-DEC-1998.

XX 02-JUN-1998; 98WO-CA00544.

XX 08-MAY-1998; 98US-0074658.

PR 03-JUN-1997; 97US-0867941.

XX (CONN-) CONNAUGHT LAB LTD.

XX Du R, Klein MH, Loosmore SM, Wang Q, Yang Y;

PI WPI: 1999-070266/06.

DR P-PSDB; AAM89413, AAM89414, AAM89415, AAM89416.

XX Lactoferrin receptor genes from Moraxella, especially M. catarrhalis

PT - useful to diagnose Moraxella infection e.g. to detect otitis media

PT due to M. catarrhalis infection and to immunise against such

PT infections

XX Claim 8; Fig 2A-2P; 202pp; English.

PS This polynucleotide comprises the lactoferrin receptor (lfr) locus

XX of Moraxella catarrhalis (Branhamella catarrhalis) 4223. There are

CC 3 tandem genes in locus, identified as lbpA, lbpB (alternative

CC start codons) and orf3A, respectively encoding lactoferrin binding

CC protein 2 (lbp2, see AAM89413), lactoferrin binding protein 1 (lbp1, see AAM89414 and AAM89415) and open reading frame protein 3 (orf3, see AAM89416). The lfr locus was identified following generation of a M. catarrhalis strain 4223 genomic DNA library and screening with a specific hybridisation probes. The genes and DNA sequences of the lfr locus are useful for diagnosis, immunisation, and the generation of diagnostic and immunological reagents. Recombinant compositions, including vaccines, based upon expressed recombinant lbp1 and/or lbp2 and/or orf3, portions of these or their analogues, can be prepared for prevention of diseases caused by Moraxella. M. catarrhalis is a causative agent of otitis media and has been associated with sinusitis, conjunctivitis and inflammatory diseases of the lower respiratory tract, such as pneumonia, chronic bronchitis, tracheitis and emphysema.

SO Sequence 7650 BP: 2410 A; 1741 C; 1636 G; 1862 T; 1 other:

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
Score:	178.00	171	118	305	324
Percent Similarity:	31.48%				
Best Local Similarity:	18.63%				
Query Match:	4,518				
DB:	20				42

US-09-494-297-2 (1-757) x AAV82019 (1-7650)

QY 4 ThrArgPheProAspLysLeuAsnThrLeuAsnThrGlnArgValLeuSerLysAsnSer 23

DB 375 ACAACTGACCCAAATGGCGATACACCACTGACACA----- 413

QY 24 LysArgPheThrValThrLeuValGlyValPheLeuMetIlePheAlaLeuValThrSer 43

DB 414 GCACAAAGACCGCGCGTCCCGCAGGTTTGTGATG----- 452

QY 44 MetValGlyAlaLysThrValPheGlyLeuValGlySerSerThrProAsnAlaLeuAsn 63

DB 453 -----GTTAAATTTCGTGATACCAAGCCCAAAAT--GAC 485

QY 64 ProAspSerSerSerGlu-----TyrArgTPTyGly-----TyrGlySerTyrVal 79

DB 486 CCAGATTATAGCAATGATTTAGTACGACGAGGCAAGCAATATATATGTTGTTATGAT 545

QY 80 ArgGlyHisPro----- 83

DB 546 GCCCATCGCCCGATGCGACAGTAAACCTTCCGTCACCCCATCCGCCCAAT 605

QY 84 -----TyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeu 98

DB 606 GACATCAAAACCTTGTATTTTAAACAARTCCCTGCATGTGTGATTCGATTTAGACAGT 665

QY 99 GluGlySerArgSerTyrGlnValTyrTyrPheAsnLeuLysLysAlaPheProLeuGly 118

DB 666 GAACGCCACCGT--TTTGACCCCAAAAAGCTAAACACCATTAAGTATGTTATGATGAC 722

QY 119 SerAspSerSerValLysLys-----TrpTyrLysHisAspGlyIleSerThrLys 136

DB 723 AACTTAACAACACCCCTTAACAAACACTTTCATCATCATCAGACAGCTATATTAAG 782

QY 137 PheGluAspTyrAlaMetSerPro-----ArgIleThr 147

DB 783 AAAAATTAACAAGCCCTGTGACCCCTTATGAAATATCCGTTGGTATCTTGACTACAA 842

QY 148 GlyAspGluLeuAsnGlnLys----- 154

DB 843 GGAAGCAGCTGACCCCAAAAATGCCGATCTCCAAATGACAAAGACCGCATTCACAA 902

QY 155 --LeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGly 173

DB 903 CCCATGCCCATTTTGTTTTATACCGA-----GAAACGCCAGCAGCAGTCCCACT 956

QY 174 LeuGluProLeuAsn----- 178

DB 957 GCTGTAAATTTAATACACAGCACTGCTGTAAGTACGATGCAAAAAGCCCT 1016

QY 179 AlaIleArgValThrGlnGluAlaValIrrpyTyrSerAspAsnAlaProIleSerAsn 198

DB 1017 GCACCTTCAGCATCGATGATGAGTGGGCTCTATCTCAATGACGAGGCAAAATCCAT 1076

QY 199 ProAspGluSerPheLysArgGlySerGlySerAsnLeuValSerThrSerGlnLeuSer 218

DB 1077 -----GAGGCGCATGTCGTCAGTCCGCCCACTTTAT 1109

QY 219 Leu----- 219

DB 1110 CTAAGCGCTTTCATATATAGACACAGCCCTGCATTCAGAGTGTATTTGACACAAAC 1169

QY 220 --MetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLys 238

```

Db      1170 TCATTACAGCAGCAGCTGTTATTATGACATCCCAAGCAGCAACTGCCCAAGCAGCAAA 1229
QY      239 GlnValProaspasphe-----GlnLeuSerIlePheGluSer 251
Db      1230 TACATCAAAAGCCATTGTGACACTACCAAAAAAGTCATGTAACCGATGTGTATCAAAATT 1289
QY      252 GlnAspLys-----GlnAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGly 269
Db      1290 GATGCCAAATTCAAACGGCAGCCGCTGTCGTGACGCCAAATCTTGTTAATGAGAAC 1349
QY      270 Leu-----ValProThrLysProProThrProGlyAspPro 281
Db      1350 ACAGAAACCCGACCTTTTATCAAGAGACCTGTTCCAAAAAGCCAAATCCCAATAAACCA 1409
QY      282 PromerProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIleGly 301
Db      1410 -----AACCTTAAT----- 1418
QY      302 AspTyrSerLysLeuLeuGlnGlyAlaThrLeuGlnLeuThrGlnAspAsnValAsnSer 321
Db      1419 -----TCAGACACCGCAGAGAGGGGCTTTATGCTAGTCGGCGCATGAGCTG----- 1466
QY      322 PheGlnAlaArgValPheSerSerAsnAsp-----IleGlyGluArg 335
Db      1467 -----GCGGGTAATTTTATCCAAATGACACACGATCTATGCTGTGGGTGAGTAAA 1520
QY      336 IleGlnLeuSerAspGlyThrTyrThrLeuThrLeuLeuAsnSerProAlaGlyTyrSer 355
Db      1521 CGAGACAAAACAGCAAACTGTCGCCCAAAAAGGTGATTTTACTGCGAGGCTTGAA 1580
QY      356 IleAlaGlnProIleThrPheLysValGlnAlaGlyLysValTyrThrIleLeuAspGly 375
Db      1581 -----AAACCTAGCAGCAGTTTGTGATATGAACGATGTGCACAAATATATACAGC 1634
QY      376 LysGlnIleGlnAsnProAsnLysGlnIleValGlu-----ProTyrSerVal 391
Db      1635 AAAAAGTTAATATGATGCGGTGAATGAGAAAATGATGATATCTTACACAGTAT 1694
QY      392 GlnAlaTyrAsnAspPhe-----GlnGlnPheSerValLeuThrThr 405
Db      1695 GAACCGCATGATGATATTTCTTGGCGCGCAAAAAAGCAGATTCACCAAAAAAGTCAGC 1754
QY      406 GlnAsn-----TyrAlaLysPheTyrTyrAlaLys 415
Db      1755 AGCAGACCCAGCCGCGCAGCTATTTTGGGCAACATGATTAATTTATTT----- 1808
QY      416 AsnLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAspLeuLysSerProPro 435
Db      1809 -----AATGGCAACTATTATGACCTATCAGCAGCAGTGTGTAATTTGGCCCTGCC 1862
QY      436 AspSerGlnAspGlyLysThrMetThrProAspPheThrIleGlyVal----- 453
Db      1863 GATGCTGTCAAAAGCCAAACCATTCATTAAGAAAAAATACCTTAATGCCACATAAATTAAG 1922
QY      453 ----- 453
Db      1923 GACAAACCAAGTACCGGCATCGTGTACAGAAAGCCAAAGATAATAGCCTTAACGCC 1982
QY      454 -----LysTyrThrIleIleAla-----GlyArgAspLeuPhe----- 464
Db      1983 ATTGCTGCCAAAAGCATGACACATCAGTTTGGCGAGCAGCGCTATAGATGCCAAAC 2042
QY      465 -----LysTyrThrValLysProArgAspPheThrProAspThrPheLeu 479
Db      2043 CAAACCCCAACAGCAGCTATTGTCAGAGCGGTAGGCGAGTACACAGCACCGCTG 2102
QY      480 LysHisIleLysLysValIle----- 486
Db      2103 CCGAAGGCGAGTAAATTCACCTTACAAAGCGCTTTGGCAGCGATCTTATCCAAAAAAG 2162
QY      487 GlnLysGlyTyrArgGlnLysGlyGlnAlaIleGlyTyrSerGly----- 501

```

```

Db      2163 GACAAAGTTATACCAATTAATGAGAAACCATCAAGAAAGGCCATCAAGATTATCTG 2222
QY      502 LeuThrGlnThrGlnLeuArgAlaIleThrGlnLeuAlaIleTyrTyrPheThr----- 519
Db      2223 TTAAACCGAAGC----- 2246
QY      520 -----AspSerAlaGlnLeuAspLysAspLysLeuLysAspTyrHisGlyPhe 535
Db      2247 GATGATGACGATGATTGACCGCATCTGATGATCAACAGATGATGATGACATGCGCAT 2306
QY      536 GlyAspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGluTyrAlaGlnAsp 555
Db      2307 GATGATTTG-----ATTGCATCTGATGATTCACAAAGAT 2339
QY      556 SerAsnProProGlnLeuThrAspLeuAspPhePheIleProAsnAsnAsnLysTyrGln 575
Db      2340 GATCACCAGATGCGCATGACAGATTCAGATGATTTG-----GGTGATGGTCGAGATGAC 2393
QY      576 SerLeuIleGlyThrGlnThrPheLysProGlnAspLeuValAspIleIleArgMetGluAsp 595
Db      2394 GCGCGCGCAGCAAAAGTATCATGACAGTAAAT-----ATTGCGCGCTGAAATT 2441
QY      596 LysLysGlnValIleProVal-----ThrHisAsnLeuThrLeu----- 608
Db      2442 GAAACCAAAATACCTGCCATTAATGAGCCTACTCATCAAAAAACCTTGCCCTAGATGT 2501
QY      609 -----ArgLysThrValThrGlyLeuAlaGly 617
Db      2502 AAAAATTAAGCTAAGTTGATGCTGATTTGACACCAACAGCCCTAAGCTGTAATTAAC 2561
QY      618 AspArgThrLysAspPheHisPheGlnIleGlnLeuLysAsnAsnLysGlnGlnLeuLeu 637
Db      2562 GATCAGAGAGGTGATATCGTCTTGATATC-----AAAATGCAAAATATGATGGCACA 2615
QY      638 SerGlnThrValLysThrAspLysThrAsnLeuGlnPheLysAspGlyLysAlaThrIle 657
Db      2616 GCGCTTACCGCCAGCCGATGTGCCAAACTATGTCGAAAGATGGGT----- 2663
QY      658 AsnLeuLysHisGlyLysLeuSerLeuThrLeuGlnGlyLeuProGlnGlyTyrSerTyrLeu 677
Db      2664 AACACCAAGGTGC-----GGTTCCTATATACAC 2693
QY      678 ValLysGlnThrAspSerGlnGlyTyrLysValLysValAsnSerGlnGlnValAla--- 696
Db      2694 ATCAAAAGATATGATGTCAAGGGCGCAATTTTGGCACAATATGCGCAAGAGTTGGCAGGG 2753
QY      697 -----AsnAlaThrValSerLysThrGly 704
Db      2754 CAGTTACAGTACGACAAAGCGATGCGATCATGATGACCGCGCAAAAAAGCAGGG 2807

```

RESULT 7
 AAF28541 standard; DNA; 49617 BP.
 AAF28541:
 04-Apr-2001 (first entry)
 Genomic fragment #28.
 Genomic library; bacteria; human upper airway; otitis media; sinusitis;
 bronchopulmonary; endocarditis; meningitis; ss.
 Moraxella catarrhalis.
 W0200078968-A2.
 28-Dec-2000.
 16-JUN-2000; 2000MO-US16649.
 18-JUN-1999; 9905-0140121.

PA (INCY-) INCYTE GENOMICS INC.
 XX Lagace RE, Patterson C, Berg KL;
 XX WPI: 2001-041427/05.
 XX
 PT Genomic library for identifying diagnostic and therapeutic
 PT compositions, and for identifying virulence factors, regulatory
 PT elements and drug targets, comprises Moraxella catarrhalis nucleic
 PT acids -
 XX
 PS Claim 1: Page 235-247; 545pp: English.
 XX
 CC The present invention relates to a Moraxella catarrhalis genomic library
 CC comprising of a combination of 41 nucleic acid molecules (see
 CC AAF28514-AAF28534). The library has a number of uses described in the
 CC specification e.g. is useful for identifying diagnostic and therapeutic
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
 CC aerobic, gram-negative diplococcus, normally found among the bacterial
 CC flora of human upper airways. M. catarrhalis is known to cause acute,
 CC localised infections such as otitis media, sinusitis and bronchopulmonary
 CC infection and life-threatening, systemic diseases including endocarditis
 CC and meningitis.
 XX
 SQ Sequence 49617 BP; 14572 A; 9853 C; 11026 G; 14166 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.0106 Length: 49617
 Score: 177.00 Matches: 164
 Percent Similarity: 33.58% Conservative: 104
 Best Local Similarity: 20.55% Mismatches: 262
 Query Match: 4.49% Indels: 268
 DB: 22 Gaps: 40
 US-09-494-297-2 (1-757) x AAF28541 (1-49617)
 OY 84 TyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGlySerArgSer 103
 DB 14887 TATTTTGATTAATTTCCCAAAATATCCGATCTG-----CACCTGAAACAGCCAG--- 14937
 OY 104 TyrGlnValTyrCysPheAsnLeuLysLysAla-----PheProLeuGly 118
 DB 14938 -----CATGTGTTGATGCTTAATAAAGCAATAATACATCAAAATATATGTTATGTT 14988
 OY 119 SerAspSerSerValLysLys-----TPTTyrLysLysHisAspGlyIle 133
 DB 14989 GCATTGTGATCACTGCGCAAAACCACTACATGAAATTATACACAAGAACCAAAACATC 15048
 OY 134 SerThrLys-----PheGlnAspTyrAla-----MetSerProArg 145
 DB 15049 AAAAAACAAAAACAGCGGATGATTTTCAAAAACATTCGTTTGGCTATATGAGCTAAGA 15108
 OY 146 IleThr-----GlyAspGlnLeuAsnGlnLysLeuArgAlaValMet 159
 DB 15109 GAGCTGAGACTTAATAAAAAAGGTGACAGACCCAGAGGACAGAACGCGTCGCATCATTT 15168
 OY 160 TyrAsnGly-----HisProGlnAsnAlaAsnGlyIleMetLeuLysLeu 174
 DB 15169 TTCACACACCTACTTATTTATTTATCATGTGTGAGATGCCAGCACCATCTGCCA----- 15222
 OY 175 GlnProLeuAsnAlaIleArgValThrGlnGlnAlaValTPTTyrTyrSerAspAsnAla 194
 DB 15223 -----AAGCGGATTAATTTGACTATGAGGGCAATGCTTATCTATGACCATGTC 15273
 OY 195 Pro-----IleSerAsnProAspGlu-----SerPhe 203
 DB 15274 AAAAAACGCCATTTTATGATTAACAGACGATAAGTAGGCATTTATTTAACTCAACC 15333
 OY 204 LysArgGlnSerGlnSerAsnLeuValSerThrSerGlnLeuSerLeu----- 219
 DB 15334 AGAAATATCAATGAAGCGATTGGTGTGAGTGCAGACACATTATATCAACAGCTTTAA 15393
 OY 220 -----MetArgGlnAla 223

DB 15394 TATAACACACACCCGCCACTTATAGCGTGCAGCTTGTGATCAAAATACCTTAACAGCAAA 15453
 OY 224 LeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAsp 243
 DB 15454 TTGTCTTATTTATGACAAACCCCAAAAGCAAGCAAGCGATGGCGTTATATAGAGTGCAG 15513
 OY 244 Phe-----GlnLeuSerIlePheGlnSerGlnAspLys----- 254
 DB 15514 TTTGATACCGACAAAAGGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 15573
 OY 255 GlyAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGlyLeu----- 270
 DB 15574 GGCACACCGCTTACTGCGCACACGCCCAAAATCTTGTATGATGATTAACACCAATACCCACT 15633
 OY 271 -----ValProThrLysProProThrProGlyAspProPheProAsn 286
 DB 15634 TTTGTTAAAGACTCTTCTCCAAAAGCAATCCCAACAAACCA-----GACCCCAAC 15687
 OY 287 GlnProGlnThrSerValLeuLeuArgLysTyrAlaIleGlyAspTyrSerLysLeu 306
 DB 15688 -----TCAGATACG 15696
 OY 307 LeuGlnGlyAlaThrLeuGlnLeuThrGlnLysAspAsnValAsnSerPheGlnAlaArgVal 326
 DB 15697 CTAGAAGCGCGGTTTATGCTGATGCTGCGGCGATGAGCTG-----CGCGGTAA 15744
 OY 327 PheSerSerAsnAspIle-----GlyGlnArgIleGlnLeuSer 339
 DB 15745 TTTTATTCCAATGACAAAGCAAGCACTTTTGTGCTTTTGGTGGCAACGACAAACGAC 15804
 OY 340 AspGlyThrTyrThrLeuThrGlnLeuAsnSerProAlaGlyTyrSerIleAlaLeuPro 359
 DB 15805 GAACCTGTGCGCAAAAAGCGGTGATTTTAACT--ACAGATTTTGAA-----AAACCC 15855
 OY 360 IleThrPheLysValGlnAlaGlyLysValTyrThrIleLeuAspGlyLys----- 376
 DB 15856 AGCACACGCTTGTGTCGATCAATGACATGATGATGATGATGATGATGATGATGATGATGAT 15915
 OY 377 ---GlnIleGlnAsnProAsnLysGlnIleValGlnProTyrSerValGlnAlaTyrAsn 395
 DB 15916 GATGAAACCAATTAATCAAAATGAAATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 15975
 OY 396 AspPhe-----GlnGlnPheSerValLeuThrThr 405
 DB 15976 GAATATTAATTATGACAGCAACCAACAAACATTCACCAAAAAATAACGCCAGCTCCAA 16035
 OY 406 GlnAsnTyrAla-----LysPheTyrTyr----- 413
 DB 16036 AAAAAACCTGCTTATTTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 16095
 OY 414 -----AlaLys-----AsnLysAsnGlySerSerGlnValValTyrCysPheAsnAla 429
 DB 16096 TTATCAGCCCAAGAAAGCAACCAAGCTTGCTGCTCCAAAGATACCAAGCAACCAATTAAGT 16155
 OY 430 AspLeuLysSerProProAspSerGlnAspGlyGlyLysThrMetThrProAspPheThr 449
 DB 16156 ATTTTGGCTAAATACCCAGATGCC-----AAAGTAAGCAGACATTAAGTT 16203
 OY 450 Thr----- 450
 DB 16204 ACCAAATTCGTTCTACAAACCAACCAAGTAAGCCGTATACCGCATTCATGCCAAAC 16263
 OY 451 -----GlyGlnValLysTyrThrHisIleAlaGlyArgAspLeuPhe 464
 DB 16264 TATGACCAACATCAGTTTGTGTAAGTATGTTATGATTAACAAAGCAACCAACACAGCG 16323
 OY 465 LysTyrThrValLysProArgAspThrAspProAspThrPheLeuLys----- 480
 DB 16324 AGTTATTTTGTCAAGCGCGTCAAGCGGATGTCACTACTCACTGCCAGTGCAGGTAAA 16383
 OY 481 -----HisIleLysValIleGlnLysGlyTyrArg 491

Db 595 GCATTGTCATCACCCTGCCAATAAACCAACCTACATGATTTATCAACAAGAACAAACATC 654
 QY 134 SerThrLys-----PheGluAspTyrAla-----MetSerProArg 145
 Db 655 AAAAAACCAAAACCAGCGATGATTCACAAAACATTCGTTTGGCTATATGAGACTAGA 714
 QY 146 IleThr-----GlyAspGluLeuAsnGlnIleuArgAlaValMet 159
 Db 715 GAGCTGACCTTAATAAAAAAGTCGACACCCAGACGACAGAACGACAAACCGTGGCATCTT 774
 QY 160 TyrAsnGly-----HisProGlnAsnAlaAsnGlyIleMetGluGlyLeu 174
 Db 775 TTCACCAACACTACTATTATTTATCATGAGAGATGCGACACCACTCTGCCA----- 828
 QY 175 GluProLeuAsnAlaIleArgValThrGlnAlaValTyrTyrTyrSerAspAsnAla 194
 Db 829 -----AAGGGGGTAAATTTGACTATGAGGCAATGGTGTATCTGACGATGTC 879
 QY 195 Pro-----IleSerAsnProAspGlu-----SerPhe 203
 Db 880 AAAAAACGCCCATTTTATGATTAACAGACGATTAAGTACGACTTATTTTAACTCAACC 939
 QY 204 LysArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeu----- 219
 Db 940 AGAAATCAATGAAGCGATTTGGTGACGACACACATTTATCTAAACGCTTAA 999
 QY 220 -----MetArgGlnAla 223
 Db 1000 TATTAACACACCCCGCCACTTATAGCGTGCCTTGTATCAAAATACCTTAAAGGCMA 1059
 QY 224 LeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAsp 243
 Db 1060 TTGCTTATATATGACAAACCCAAACACAAACACCCGATGGCGTATATATGAGAATCAG 1119
 QY 244 Phe-----GlnLeuSerIlePheGluSerGluAspLys----- 254
 Db 1120 TTTCATACCGCAAAAAGTCATGAAGCCGATGCTATGATGATGACGCCAAGATTAA 1179
 QY 255 GlyAspLysTyrAsnLysGlyTyrGlnAsnLeuSerGlyGlyLeu----- 270
 Db 1180 GGCACACGCTTACTGCGACAGCCAAATCTTGTATGATGATTAACACCAATACCGCACT 1239
 QY 271 -----ValProThrLysProProThrProGlyAspProPrometProAsn 286
 Db 1240 TTTCGTTAAAGAGCTTCTCCAAAACCAATCCCAACACCA-----GACCCCAAC 1293
 QY 287 GlnProGlnThrSerValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeu 306
 Db 1294 -----TCAAGATACG 1302
 QY 307 LeuGlnGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgVal 326
 Db 1303 CTAAAGACGGGTTTATGTGTAGTCGGGATGAGCTG-----GCGGGTAA 1350
 QY 327 PheSerSerAsnAspIle-----GlyGluArgIleGluLeuSer 339
 Db 1351 TTTTATCAATGACAAACGCAACTTTTGTGCTCTTGGTCAACAGACAAACGAC 1410
 QY 340 AspGlyThrTyrThrLeuThrGluLeuAsnSerProAlaGlyTyrSerIleAlaGluPro 359
 Db 1411 GAACCTGCGCCACAAAACGGTGTATTTAGT---ACAGATTTGA-----AAACC 1461
 QY 360 IleThrPheLysValGluAlaGlyLysValTyrThrIleIleAspGlyLys----- 376
 Db 1462 AGCCACGAGCTTGTGGCAATGAAGATGTGTCATTTATTTGCGTAAAAAAGTTAA 1521
 QY 377 ---GlnIleGluAsnProAsnLysGluIleValGluProTyrSerValGluAlaTyrAsn 395
 Db 1522 GATGAAGTCATATCAATCAATTTGAAGTAAACTGCTCCCTGATTAATAAGATTTAT 1581
 QY 396 AspPhe-----GluGluPheSerValLeuThr 405
 Db 1582 GAATATATATATGACAGACCAACAAACATTCACAAAAAATAAACGCCAGCGTCCAA 1641

QY 406 GlnAsnTyrAla-----LysPheTyrTyr----- 413
 Db 1442 AAAAACCTGCTTATTTTGGTACAGATGATGATTTTATTTATGATGATTAATGAC 1701
 QY 414 -----AlaLys-----AsnLysAsnLysSerGlnValValTyrCysPheAsnAla 429
 Db 1702 TTATCAGCCAAAGAAAGCAAAACAGCTTGCTCTCCCAAGTATACCAGACCAATTAAGAGT 1761
 QY 430 AspLeuLysSerProProAspSerGluAspGlyLysTyrThrMetThrProAspPheThr 449
 Db 1762 ATTTTGGCTAAATATCCAGATGCC-----AAAGTACACACAGCAATTAAGTT 1809
 QY 450 Thr----- 450
 Db 1810 ACCAAATCGTTTCAACAACAGCAAGATTAAGCCGATATACGCCCATTCATCCAAAAAGC 1869
 QY 451 -----GlyGluValLysTyrThrHisIleAlaGlyArgAspLeuPhe 464
 Db 1870 TATGACACATCAGTTTGGTGAAGTATTTGATATATGATTAACAAGGCAACCCACACGC 1929
 QY 465 LysTyrThrValLysProArgAspThrAspProAspThrPheLeuLys----- 480
 Db 1930 AGTATATTTTGTCAAGCGGTCAAGCGATGCTACTGACTGACGCGCCAGTCAAGTAA 1989
 QY 481 -----HisIleLysValIleGluLysGlyTyrArg 491
 Db 1990 TTCACCTAATATGCTTTTGGGACGAGCTGACCCGAAAAAGCAAAAGCTTAATAGC 2049
 QY 492 GluLysGlyGlnAlaIleGluTyrSerGlyLeuThrGlnLeuArgAlaIleThr 511
 Db 2050 AAAGTACAGATACCATCAACCAAAAGCTTTAAAGAT----- 2088
 QY 512 GlnLeuAlaIleTyrTyrPheThr----- 519
 Db 2089 -----TATATATGACCAAGACTTTATCCCAAGATGACGATGACATGAC 2136
 QY 520 AspSerAlaGluLeuAspLysAspLysLeuLysAsp---TyrHisGlyPheGlyAspMet 538
 Db 2137 GATGATTTGACCGCATGATGATGATCAACAGATGATATATACATGCGGATGATGATTTG 2196
 QY 539 AsnAspSerThrLeuAlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnPro 558
 Db 2197 -----ATTCACTGATGATTTACAGATGATGACACA 2229
 QY 559 ProGlnLeuThrAspLeuAspPhePheIleProAsnAsnAsnLysTyrGlnSerLeuIle 578
 Db 2230 GATGCGCATGACGATTCAGATGATTTG-----GGTATGGTGCAGATGATGACGCCCA 2283
 QY 579 GlyThrGlnTyrPheIleProGluAspLeuValAspIleIleArgMetGluAspLysGlu 598
 Db 2284 GGCAAAGTGTATCATGACGATTAAT-----ATTCGCTCCGATTTGAAAACAAA 2331
 QY 599 ValIleProVal-----ThrHisAsnLeuThrLeu----- 608
 Db 2332 TACTTGCCATTAAATGACCTTACTCATGAATAAAACCTTTGCCCTAGATGTAATAATAG 2391
 QY 609 -----ArgLysThrValThrGlyLeuAlaGlyAspArgThr 620
 Db 2392 GCTAAGTTTGAATGAATCTTGACACCAACAGCCTAAGTGTAAATTAACGATGACAGA 2451
 QY 621 LysAspPheHisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThr 640
 Db 2452 GGTATATTCGCTTTGATATC-----AAAATGCAAAAATTTGATGCGACAGATTTAAC 2505
 QY 641 ValLysThrAspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLys 660
 Db 2506 GCCAAAGCGGTGCGCAACACTATCGTGAAGAAAGTGGT-----AACACCAA 2553
 QY 661 HisGlyGluSerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGlu 680
 Db 2554 GGTGGC-----GGTTCTTATACACATCAAGAT 2583

QY 681 ThrAspSerGluGlyTyrLysValLysValAsnSerGlnIValAla 696
||| ::|||
DB 2584 ATTGATGTTAAGGGCAATTTTGGCACAAATGGCAAGATTGGCA 2631

RESULT 9
AB069442
ID AB069442 standard; DNA; 4161 BP.
XX
AC AB069442:
DT 29-AUG-2002 (first entry)
XX
DE Listeria innocua DNA sequence #881.
XX
KM Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
OS Listeria innocua.
XX
PN W0200228891-A2.
XX
PD 11-APR-2002.
XX
PF 04-OCT-2001; 2001WO-FR03061.
XX
PR 04-OCT-2000; 2000FR-0012697.
XX
PA (INSP) INSP PASTEUR.
XX (CNRS) CNRS CENT NAT RECH SCI.
PI Kunst F, Glaser P.
XX
DR WPI; 2002-332479/37.
XX
PT New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators
PS
PS Claim 7: SEQ ID 2255; 180bp; French.
XX
XX The present invention relates to nucleic acid sequences
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be
CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in
CC anti-Listeria vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SO Sequence 4161 BP; 1660 A; 657 C; 815 G; 1029 T; 0 other;

Alignment Scores:
Pred. No.: 0.00888 Length: 4161
Score: 161.00 Matches: 163
Percent Similarity: 34.98% Conservative: 127
Best Local Similarity: 19.66% Mismatches: 301
Query Match: 4.08% Indels: 238
DB: 24 Gaps: 41

US-09-494-297-2 (1-757) x AB069442 (1-4161)

QY 10 LeuAsnThrLeuAsnThrGlnArgValLleuSerLysAsnSerLys---ArgPheThrVal 28
||| ::|||
DB 988 CTCAGTAGCTTTCATCAAAAAGATGACATATATGACAAAGCATGTTTATGATATTCGAA 1047

QY 29 ThrLeuValGlyValAlaPheLeuMetLlePheAlaLeuValThrSerMetValGlyAlaLys 48
::|||
DB 1048 AGCTGACAAAGCTTATGATATAGCAGTCTTGAAGCAATAGTATGCTACAACTGAATGAAG 1107

QY 49 ThrValPhe-----GlyLeuValGlu----- 55
||| ::|||
DB 1108 TTATGTTTACCGAGGACCAATATTTGTCCTCATTTAAGTTAGCAAAAAAATAAGTTT 1167

QY 56 -----SerSerThrProAsnAlaIleAsnProAspSerSerSerGluTyrArg--- 71
::|||
DB 1168 TATACGTGGCTACGATGAAATCAAACTGTTAATCCACCCCGCAGTCACCTTATTAATGCA 1227

QY 72 ---TrpTyrGlyTyrGluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgVal 90
|||
DB 1228 ACGTGG-----ACGACCCCTCGTACAAACCAACAAATTTGG 1263

QY 91 AlaHisAspLeuArgValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsn 110
::|||
DB 1264 TCAGACTCTTTAAGCATGAAATGACAGCGGACCATGATGATATATATCATTTGGTCT 1323

QY 111 LeuLysLysAlaPheProLeuGlySerAspSerSerVal-----LysLysTrpTyr 127
::|||
DB 1324 GCTGTAAATTAATATC---ATAGCCATTTGATCTACATTTATTCAGGTGATGCTGGAA 1380

QY 128 LysLysHisAspGlyLysSerThrLysPheGluAspTyrAlaMetSerPro----- 144
::|||
DB 1381 GCTGAGGATTAATTTATTTTCAGCAGCAGATTAAGATTTCAATCTATGATTTCCGCTG 1440

QY 145 ---ArgLleThrGly-----AspGluLeuAsnGlnLysLeuArgAlaValMetTyrAsn 161
::|||
DB 1441 GTACAGATTAACAGATCCGATACGATACGACACCTGGTGAATATGAAATACATATATAG 1500

QY 162 GlyHisProGlnAsnAlaAsnGlyLleMetGlyGlyLeuGluProLeuAsnAlaIleArg 181
::|||
DB 1501 -----GTAAATGGCTTACATCTACATCAATCACACGATCA 1533

QY 182 ValThrGlnGlnAlaValTrpTyrTyrSerAspAsnAlaProLleSerAsnProAspGlu 201
||| ::|||
DB 1534 GTCAAGAAATTAATACCAAGTCTTTAGCGGAAATCACTACTTCTATCA---AAAGAA 1590

QY 202 SerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuMetArg 221
::|||
DB 1591 TCCTGGAAA-----GCCGAGATTAATTTGTTTCAGCAACCAAT-----AAAAA 1635

QY 222 GlnAlaLeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValPro 241
::|||
DB 1636 GGTAAATTCAGATATTTTCGTCAGTTACCGTACAGGGGAGATGATGTAATATACACA 1695

QY 242 AspAspPheGlnLeu----- 246
||| ::|||
DB 1696 GAGAGCTACCAATTTATGTTATACATTTGATGAGCTCTACAAATAATTTGACAGGTG 1755

QY 247 -----SerLlePheGluSerGluAspLys-----GlyAspLysTyr 258
||| ::|||
DB 1756 AAGAGATTAAGACGATATAGCAAGCTAAAGATTCATTTCTATATAGTGTATACGTGG 1815

QY 259 AsnLys-----GlyTyrGln 263
|||
DB 1816 AATTCGAAGACATTTTATTTTCAGCAGACAGATTAAGATGTATTCGGGTGACCTTAAG 1875

QY 264 AsnLeuLeuSerGlyLysLeuValProThr---LysProProThrProGlyAspPropo 282
::|||
DB 1876 GATATTAAGTACAGACGACAGTAAACAGATTAACAGCAAGCAACCAAGTACGTAT 1935

QY 283 MetProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIleGlyAsp 302
::|||
DB 1936 TTGTATGAAACCAAGTAAAGCAAGTACATTAACGATCAAA-----CGCGAT 1983

QY 303 TyrSerLysLeuLeuGlnGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPhe 322
|||
DB 1984 CAATCAACTTTAAGGCCCAAGATTCATATTTATATGCGAATAGTGAAGTGCACAA 2043

QY 323 GlnAlaArgValPheSerSerAsnAspLleGlyLysArgLleGluLeuSerAsp----- 340
::|||
DB 2044 GACATTTTATTAACGCCACAGTAAAGATGTAATCCAGTACATTTAAGATATATGAA 2103

QY 341 -----GlyThrTyrThrLeuThrGluLeuAsnSerProLagLysTyrSerLleAlaGlu 358

```

Db      2104 GTCCAGGCTACGACAGACACA-----AAA 2130
Oy      359 ProIleThrPheIysValGluAlaGlyValIleThrIleIleAspGlyLysGlnIle 378
Db      2131 CCAGCACTTAAACAAGATT-----ACGTATATATATGA----- 2163
Oy      379 GluAsnProAsnLysGluIle-----ValGluProTyrSerValGluAla 393
Db      2164 ---AATCTAAGTAAGAAAGATTACCGTAACAGTCAAGCAAGCAACTTTAGAACCC 2220
Oy      394 TyrAsnAspPheGluGluPheSerValIleThrGlnAsnTyrAlaLysPheTyrTyr 413
Db      2221 ---AAGATTCAGCTCTTATGATAGTGATACGCGAATCCAAAGATATTTATTTCA 2277
Oy      414 AAlaLysAsnLysAsnGlySerSerGlnValValIleTyrLysPheAsnAlaAspLeuLysSer 433
Db      2278 GCAACCGCATTAAGATGGC-----ACT 2298
Oy      434 ProProAspSerGluAspGlyGlyLysThrMetThrProAspPheThr---ThrGlyGlu 452
Db      2299 CCACTAGACTTTAAGATATATAAAGTCGACGACAGTAGACACACGCAAGCAAGCAAGCAACT 2358
Oy      453 ValLysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProArgAsp 472
Db      2359 AACAGGTAACCTATTGTGTATGAACCAAGTAAGCAAGTACCGTAAGCTCAAA--- 2415
Oy      473 ThrAspProAspThrPheLeuLysHisIleLysLysValIleGluLysGlyTyrArgGlu 492
Db      2416 GCGGATCAACACACTTA-----GAG 2436
Oy      493 LysGlyGlnAlaIleGluTyrSerGly-----LeuThrGluThrGluLeuArgAla 509
Db      2437 CTCAAAGATTTCAGTATTATTAATGACGACAAAGTGAAGCAAGCAAGTATTATTTCAC 2496
Oy      510 AlaThrGlnLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLysAspLysLeu 529
Db      2497 GCCACA-----GACCAAAACC 2511
Oy      530 LysAspTyrHisGlyPheGlyAspMetAsn-----AspSerThrLeuAla 544
Db      2512 GGAATTCAGTACGCTTAAAGCATTAATGAAGTCGACAGTACAGTACACAAACGAAAC 2571
Oy      545 ---ValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProGluLeuThr--- 562
Db      2572 GGAACATAACAGGTTACCTATACGATGCAATCAACAAAGCAAGTACATACACATGC 2631
Oy      563 -----AspLeuAspPhePheIleProAsnAsnAsnLysTyrGlnSerLeu 577
Db      2632 AAACCGCATCAAGCACTTTAGAA-----GCCAAAGATTCAATT 2670
Oy      578 Ile-----GlyThrGlnTyrHisProGluAspLeuValAspIleIleArgMetGluAsp 595
Db      2671 ATTATATACCGAGATTAAGTGAAGACAGAAC-----AACTTATTTCCGCTACTGAT 2724
Oy      596 LysLysGluValIleProValThrHisAsnLeuThrLeuAlaGlyLysValThrValThrGlyLeu 615
Db      2725 AAACAGAGAAAGACAATTGACTTTAAATATTTAAAGGAGAGACACATA----- 2775
Oy      616 AlaGlyAspArgThrLysAspPheHisPheGluIleGluLeu----- 629
Db      2776 -----GATACCACAAAAAGCTGTAATTATGACATCACTTTCTTATTACGAGTCACT 2829
Oy      630 ---LysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThrAspLysThrAsnLeu 648
Db      2830 CGTAGTACTGAATTGATTAACAAATTAACACTAACAGCAAAAAAACAAGTAAACTTGG 2889
Oy      649 GluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyLysLeuThrLeuLys 668
Db      2890 GAACGAAAGACTCACT----- 2907
Oy      669 GlyLeuProGluGlyTyrSerTyrLeuValLysGlu-----ThrAspSer 683

```

```

Db      2908 ---CTTATGAAGAGACAGATGATACGTAAGATTAAGTCTGTTTCAGCAACGATAAA 2964
Oy      684 GluGlyTyrLysValLysValAsnSerGlnGluValAlaAsnAla----- 698
Db      2965 GATGGCAATACAGTGACTTTAAAGCTATGTGAAGTTAAAGAACTGTAATACCAAAAA 3024
Oy      699 -----ThrValSerLysThrGlyIleThrSerAspGluThrLeuAla 712
Db      3025 GCAGCACTTACAAATATCACCATTATATGAGGATCAGGACATCAGACACATTAACAGTGACA 3084
Oy      713 PheGluAsnAsnLysGluProValVal 721
Db      3085 GTTTCAGCAATCAACAACTAAATCGTA 3111

RESULT 10
AB067871
ID      AB067871 standard; DNA; 4185 BP.
XX      XX
AC      AB067871;
XX      XX
DT      29-AUG-2002 (first entry)
XX      XX
DE      Listeria innocua DNA sequence #673.
XX      XX
KW      Antibacterial; Listeria; food contamination; mutational analysis;
        infection; ds.
OS      Listeria innocua.
XX      XX
PN      W0200228891-A2.
XX      XX
PD      11-APR-2002.
XX      XX
PF      04-OCT-2001; 2001WO-FR03061.
XX      XX
PR      04-OCT-2000; 2000FR-0012697.
XX      XX
PA      (INSP ) INST PASTEUR.
XX      XX
PI      (CNRS ) CNRS CENT NAT RECH SCI.
XX      XX
PT      Kunst F, Glaser P;
DR      WPI, 2002-332479/37.
XX      XX
PT      New genomic sequences from Listeria species, useful for detection,
        treatment and prevention of infection, also related polypeptides,
        antibodies and modulators
XX      XX
PS      Claim 7; SEQ ID 684; 180pp; French.
XX      XX
CC      The present invention relates to nucleic acid sequences
        (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
        and primers for identification and/or detection of Listeria (e.g. as
        contaminants in foods, or mutational analysis) and for analysis of
        gene expression. Proteins encoded by the nucleic acid sequences can be
        used to screen for compounds that modulate gene expression, replication
        and pathogenicity of Listeria (potential therapeutic agents), also for
        treating infections by Listeria, and are useful as immunogens in
        anti-Listeria vaccines.
CC      Note: The sequence data for this patent did not form part
        of the printed specification, but was obtained in electronic format
        directly from WPI at ftp.wipo.int/pub/published_pct_sequences.
XX      XX
SQ      Sequence 4185 BP; 1671 A; 658 C; 821 G; 1035 T; 0 other;

Alignment Scores:
Pred. No.: 0.00894      Length: 4185
Score: 161.00      Matches: 163
Percent Similarity: 34.98      Conservative: 127
Best Local Similarity: 19.66      Mismatches: 301
Query Match: 4.08      Indels: 238
DB: 24      Gaps: 41

```


US-09-494-297-2 (1-757) x ABQ67871 (1-4185)

OY	10	LeuasnThrLeuasnThrGlnArgValLeuSerLysasnSerLys---	ArgPheThrVal	28
		1012	CTCAGTACTCTTTGATCAAAAAAATGCGATATATGACACACGCTTTGATATTCGGAA	1071
OY	29	ThrLeuValGlyValPheLeuMetIlePheAlaLeuValThrSerMetValGlyAlaLys	48	
		1072	AGTCGACACAGCTTATGATATTAAGCAGCTTTGAAACGATACGTCTTACAACTGAAATGAAC	1133
OY	49	ThrValPhe---	GlyLeuValGlu---	55
		1132	TTAATGTGTGACGGACCAATTAATTTGCTCCCTACATTAAGTATAGAAAAAATAAAGTTT	1191
OY	56	-----SerSerThrProsnAlaIleasnProAsnSerSerSerGluTyrArg---	71	
		1192	TATACGTGGCTAGCTATGAAAAAACTGTTATCCACCCCGCAGCGACCTTATATGGA	1251
OY	72	---TrpTyrGlyTyrGluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgVal	90	
		1252	ACGCTG-----ACGAACCTCGTACAAACGAACAATTTTG	1287
OY	91	AlaHisAspLeuArgValAsnLeuGlnGlySerAsnSerTyrGlnValTyrCysPheasn	110	
		1288	TCAGACTTTTAAACGATGAATAATACGACGCTGCAACCATGATGATATTAATCATTTGGCT	1347
OY	111	LeuLysLysAlaPheProLeuGlySerAsnSerSerVal-----LysLysTrpTyr	127	
		1348	GCTGTAAATATATATC---ATAGCACTGATCTCTACTATTTATTCAGTGTAGCTGGAAA	1404
OY	128	LysLysHisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerPro-----	144	
		1405	GCTGAGGATATTTTATTTATTCAGCAACGAGATTAAGCTTTCATTAATGATTTTCCGCTG	1464
OY	145	---ArgIleThrGly-----AspGluLeuasnGlnLysLeuArgAlaValMetTyrasn	161	
		1465	GTAACAGTAACAGGATCCGTAGATATCTACGACACCTGTGAATATGAANAATCAATATAGT	1524
OY	162	GlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluProLeuAsnAlaIleArg	181	
		1525	-----GTAATGCGTTGACTACTACATCACAGTACA	1557
OY	182	ValThrGlnGluAlaValTrpTyrTyrSerAspAsnAlaProIleSerAsnProAspGlu	201	
		1558	GTGAAGAAGAAATCAACGACAGCTGTGACGGAANAATCACTACTACTATACA---AAAGAA	1614
OY	202	SerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuMetArg	221	
		1615	TCGGGAAA-----GCCGAGAAATATTTTCTTCAGCAACCAAT-----AAAAAA	1659
OY	222	GlnAlaLeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValPro	241	
		1660	GGTAAATTCAGAGATATTTCTCTCAGTTACCGTAACAGGGAAATTGATTAATATCACCA	1719
OY	242	AspAspPheGlnLeu-----	246	
		1720	GGAGACATACGAATATATGATATACAAATGATGAGCTCTACAAAAATATGTGACGCTG	1779
OY	247	-----SerIlePheGluSerGluAspLys-----GlyAspLysTyr	258	
		1780	AAAGAGATAAACACAGTATGAAGCTAAAGATTCATTTCTTATATAGGTACACGCGG	1839
OY	259	AsnLys-----GlyTyrGln	263	
		1840	AATTCCAAAGCAATTTTATTTACGCGACAGATTAAGATGGTAAATCCGGTGGACTTTAAG	1899
OY	264	AsnLeuLeuSerGlyGlyLeuValProThr---LysProProThrProGlyLysProPro	282	
		1900	GATATTTAAAGTAGAAGACACAGTAAACACGATTAACACAGCAACTAACAAAGTTACGAT	1959
OY	283	MetProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIleGlyAsp	302	
		1960	TTGTATGTGAACCAAGATGAAGGATTAACCTTAACACTCAAA-----GCCGAT	2007

QY	303	TyrSerLysLeuLeuGluGlyValAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPhe	322
Db	2008	CAATCAACTTATTACGAGCCCAAAAGATTCAATTTATTATCTGGGAGTAAGTGAATGCCAAA	2067
QY	323	GlnAlaArgValAPheSerSerAsnAspIleGlyGlnArgIleGluLeuSerAsp	340
Db	2068	GACAACTTTTATTACGACCAAGATTAAAGATGTGTAATCCAGTAAGCTTAAAGATATTGAA	2127
QY	341	-----GlyThrTrpThrLeuThrGluLeuAsnSerProAlaGlyTyrSerIleAlaIu	358
Db	2128	GTCGAAGGTACAGTACAGACACAACA-----AAA	2154
QY	359	ProIleThrPheLysValGlnAlaGlyLysValTyrThrIleIleAspGlyLysGlnIle	378
Db	2155	CCAGAACTATACAGAGATT-----ACGTATATATATGCA-----	2187
QY	379	GluAsnProAsnLysGluIle-----ValGluProTyrSerValGluAla	393
Db	2188	---AATCTAACTAAAGGAAGTTACCGTACAGTCAAGTCAAGCAAGCACTTTGAGAACCC	2244
QY	394	TyrAsnAspPheGluGluPheSerValLeuThrTrpGlnAsnTyrAlaLysPheTyrTrp	413
Db	2245	---AAAGATTACGCTCTTATGTAGCGTATACGTGGAAATTCCAAAGATAATTATTCTCA	2301
QY	414	AlaLysAsnLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAspLeuLysSer	433
Db	2302	GCAACCGCATAAAGATGCG-----ACT	2322
QY	434	ProProAspSerGlyAspGlyGlyLysThrMetThrProAspPheThr---ThrGlyGlu	452
Db	2323	CCAGTAGCTTTAAAGATATATAAAGTCGACAGCACAGTAGACACCCAGCAAAAGCAGGAAC	2382
QY	453	ValLysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProArgAsp	472
Db	2383	AACAAGGTAACTATTGTGTATGAAACCAAGTAAGAAAGTTACCGTAAACGTCAA---	2439
QY	473	ThrAspProAspThrPheLeuLysHisIleLysLysValIleGluLysGlyTyrArgGlu	492
Db	2440	GCGGATCAACAACTTAA-----GAG	2460
QY	493	LysGlyGlnAlaIleGlyTyrSerGly-----LeuThrGluThrGlnLeuArgAla	509
Db	2461	GTCAAAGATTACGATTATTATACGTGACACAAAGTCGAAGAAGCAAGATAACTTATTCTCA	2520
QY	510	AlaThrGlnLeuAlaIleTyrTyrTrpPheThrAspSerAlaGluLeuAspLysAspLysLeu	529
Db	2521	GCCCACA-----GACAAAAAC	2535
QY	530	LysAspTyrHisGlyPheGlyAspMetAsn-----AspSerThrLeuAla	544
Db	2536	GGAATTCAGTAGGCTTTAAAGATATTAAAGTCGACAGGTACAGTAGACACACAAAGCA	2595
QY	545	---ValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProProGlnLeuThr---	562
Db	2596	GGAACATAACAAGTTACCTATACGTATGCAAAATCAAGCAAGAAATTACAGTAAACAGTC	2655
QY	563	-----AspLeuAspPhePheIleProAsnAsnAsnLysTyrGlnSerLeu	577
Db	2656	AAAGCGGATCAACAGCACTTTAGAA-----GCCAAAGATTCAATT	2694
QY	578	Ile-----GlyThrGlnThrPheHisProGluAspLeuValAspIleIleArgMetGluAsp	595
Db	2695	ATTATATACCGAGATAGAGTGAAGAACAGAACAC---AACTTTATTTCCGCTACTGAT	2748
QY	596	LysLysGluValIleProValThrHisAsnLeuThrLeuArgLysThrValThrGlyLeu	615
Db	2749	AAACGAGAAAGACAAATGCACTTTAAATAATTTAAAGTGAAGGACAAAT---	2799
QY	616	AlaGlyAspArgThrLysAspPheHisPheGluIleGluLeu-----	629
Db	2800	-----GATACCCACAAAGAGTGTGAATATATGACATCACTTATTTCTTATTCAGAGTACT	2853

```

QY 630 ---LysAsnAsnLysGlnLeuLeuSerGlnThrValLysThrAspLysThrAsnLeu 648
Db 2854 CGTACTGACTGATTTAGTAAACAACTTACAGTACAGTCAAAAAATCACTAACTTG 2913
QY 649 GlnPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlySerLeuThrLeuGln 668
Db 2914 GAAGCGAAGAGCTCAACT-----2931
QY 669 GlyLeuProGlyLysThrSerTyrLeuValLysGlu-----ThrAspSer 683
Db 2932 ---CTTTATGAGAGGACAGCACTGATGCTTAAAGTAACTTTGTTTCAGCAACCATTA 2988
QY 684 GlnGlyTyrLysValLysValAsnSerGlnGlnValAlaAsnAla-----698
Db 2989 GATGCAATACAGTGGACTTTAAAGCTATTGAGTTAAAGCACTGTAATACCAAAA 3048
QY 699 -----ThrValSerLysThrGlyIleThrSerAspGlnThrLeuAla 712
Db 3049 GCAGCACTTACAAATAATCACCTATTCATATGACAGCATCAGCAACAAATTACACTGACA 3108
QY 713 PheGlnAsnAsnLysGlnProValVal 721
Db 3109 GTCTAGCCCAATCAAACTAAATCATGTA 3135

```

RESULT 11
AAx20201
ID AAx20201 standard; DNA; 4249 BP.

```

XX AC AAx20201;
XX DT 20-APR-1999 (first entry)
XX DE Enterococcus faecalis EF104 gene fragment.
XX KM Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX KM detection; attenuation; antigenic; ss.
XX OS Enterococcus faecalis.
XX PN W09850554-A2.
XX PD 12-NOV-1998.
XX PF 04-MAY-1998; 98WO-US08959.
XX PR 14-NOV-1997; 97US-0066009.
XX PR 06-MAY-1997; 97US-0044031.
XX PR 16-MAY-1997; 97US-0046655.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Bailey C, Choi GH, Hiromocky J A, Kunsch CA;
DR WPI; 1999-070095/06.
DR P-PSDB; AAY00211.
XX PT New isolated Enterococcus faecalis polynucleotides - used to develop
XX PT products for the detection of Enterococcus and for use in vaccines
XX PT for prevention or attenuation of Enterococcus infection
XX PS Claim 1: Page 202-203; 301pp; English.
XX CC The present sequence encodes an antigenic polypeptide fragment
XX CC isolated from Enterococcus faecalis. The present invention describes
XX CC genes, proteins and antigenic polypeptides isolated from E. faecalis.
XX CC The proteins can be used in vaccines for preventing or attenuating an
XX CC infection caused by a member of the Enterococcus genus in an animal.
XX CC They can also be used for detecting Enterococcus antibodies in a sample.
XX CC The nucleotide sequences can be used for detecting Enterococcus nucleic
XX CC acids. Products from the present invention can also be used for
XX CC screening compounds to identify agonists and antagonists of E. faecalis
XX CC protein activity.

```

```

SQ Sequence 4249 BP; 1579 A; 779 C; 815 G; 1069 T; 7 other;
Alignment Scores:
Pred. No.: 0.00911 Length: 4249
Score: 161.00 Matches: 159
Percent Similarity: 34.83% Conservative: 104
Best Local Similarity: 21.06% Mismatches: 326
Query Match: 4.08% Indels: 166
DB: 20 Gaps: 40
US-09-494-297-2 (1-757) x AAx20201 (1-4249)
QY 39 AlaLeuValThrSerMetValGlyAlaLysThrValPheGlyLeuValGluSer-----56
Db 1451 TCATTAGTACACCTGTAAATGCTCTTAATAAGCCATTCAATTAGTATCCGATCAATAT 1510
QY 57 ---SerThrProAsnAlaIleAsnProAspSerSerSerGluTyrArgTyrGlyTyr 75
Db 1511 ATTGAACCAATTAGCTGTATTATACCT---TTGAATGCTGAACACGCTGGGCTAATTAT 1567
QY 76 GluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
Db 1568 GATCAA-----AATGTCCTATTCATCAAGA 1594
QY 96 -----ValAsnLeuGlyLysArgSerTyrGlnValTyrCysPheAsnLeuLys 113
Db 1595 ACACTGCTCAGTATGAGGACAAAGCAAGCAACCGATTCAAATTTGAATTAAGTA 1654
QY 114 AlaPhePro-----LeuGlySerAspSerSerValLysLysTyrTyrLysLys 129
Db 1655 AAGCATCCTAATTATCTTCAATTACGACTACAAAGAAATTAATTTATTATACAG---1711
QY 130 HisAspGlyIleSerThrLysPheGlnAspTyrAlaMetSerProArgIleThrGlyAsp 149
Db 1712 -----TTAGGAACG-----GATTATACAGTAACGCCCAAGCTCAGATGTTCA 1753
QY 150 GluLeuAsnGlnLysLeuArgAlaValIleMetTyrAsnGlnLysProGlnAsnAlaAsnGly 169
Db 1754 GTTATTAACTTACTAGCCCAATTAACAGCAAAATTCAAATTCGATTGTTTAAATTAT 1813
QY 170 IleMetGlyLysLeuGlu-----ProLeuAsnAlaIleArgValThrGln 184
Db 1814 GTGCCAGATAGTGTTCACAAAGATAAAGTAAAGTATCCAGTCGATACGATACAGTAC 1873
QY 185 GluAlaValTyrPyrTyrSerAspAsnAlaProIleSerAsnProAspGluSerPheLys 204
Db 1874 AGTCTGAAGTTTAACTCCAGTTGATGACGACGATACACTACT-----AATAGTAA 1924
QY 205 ArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuMetArgIleAlaLeu 224
Db 1925 CGTGTCTTGACGACCACTTCAAAGTAAAGTAAATTCATTCTTGTCAATGACGAAAT 1984
QY 225 LysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAspPhe 244
Db 1985 GATTCCTTTTGACTCACTAGCGCTCGCTACAAATAATTCAGCGGCGCC-----GATGTT 2038
QY 245 GlnLeuSerIlePheGluSerGlnAspLysGlyAspLysTyrAsnLysGlyTyrGlnAsn 264
Db 2039 CTTTGTGACATTATGATGATTTCAAAAC-----GATCAGGATGATCAATTTATCCACA 2092
QY 265 LeuLeuSerGlyGlyLeuValProThrLysProProThrProGlyAspProPheMetPro 284
Db 2093 TACTGGGACCGCGGCACTACTTTGATATTAACCAATGACGCCCAACAGCCCTGGATATCA 2152
QY 285 -----ProAsnGlnProGlnThrThrSerValLeuIle-----Arg 296
Db 2153 ACGATTACTTTTGACGAAATAATCAATAGTATACAGCTTGATTTGGAAAAACCAAGAA 2212
QY 297 LysTyrAlaIleGlyAspTyrSerLysLeu-----LeuGlnGlyAlaThrLeuGln 313
Db 2213 CATTACATTATTT---GAGTATAAAAACCCCAATGCTGATGCTGCCAATCTTTAT 2269
QY 314 LeuThrGly-----AspAsnValAsnSerPheGlnAlaArgValPhe 327

```

Db	2270	ATTAACAGGACGACGAAGAAGCAACCAATCGAATTAATTAATGAAGCTCTGCTTCGGTTTCT	2325
OY	328	SerSerAsnAspIleGlyGluArgIleGluLeuSerAspGlyThrTyrThrLeuThrGlu	347
Db	2330	GTTCAAAAT-----GAAGCGTTAGACATTGTTGAGTCGAACA-----	2365
OY	348	LeuInsSerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAlaGly	367
Db	2366	-----CAAGCGGCGAATCCACATTTAAAAATGTAAACAAAACG	2404
OY	368	LysValTyrThr-----IleIleAspGlyLys-----GlnIleGluAsnProAsnLysGlu	384
Db	2405	ACAGTAACACACAAAAAATATTGATATATAAAACACATCGCTGGAAAAATCCACACGTTTAA	2466
OY	385	IleValGluProTyrSerValGluAlaTyrAsnAspPheGluGluPheSerValLeuThr	404
Db	2465	TTAACACCAAAAGGACCAACCAACGCTCAATTCGATTTCATTATACCGTG-----	2518
OY	405	ThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAsnLysAsnGlySerSerGlnValVal	424
Db	2519	-----AAAGCGGTCGCCAGAAAGATGCT	2533
OY	425	TyrCysPheAsnAlaAspLeuLysSerProProAspSerGluAspGlyLysThrMet	444
Db	2540	TAT-----TCATTAGAGAGACTACACAAAGGTCGGAATCTCAT	2578
OY	445	ThrProAspPheThrThrGlyLys-----ValLysTyrThrHisIleAlaGlyArg	461
Db	2579	TTTAAAGCTGTACTTACCTTTCAGCAAAAAACATTACGATGTGATATACATACGCTCTCTCAAC	2633
OY	462	AspLeuPheLysTyrThrValLysProArgAspThrAspProAspThrPheLeuLysHis	481
Db	2639	GCTGGCCAAATCTATACAGAAAACAACATCGACTCGTAACATTGAACACAGATGCTGCT	2698
OY	482	IleLysLysValIle-----GluLys	488
Db	2699	AGCAAGAAAAAAGTCCACCACCTGCCCAATGCACATTGAAATTCGAGAAGTGATCGGAA	2758
OY	489	GlyTyrArgGluLysGlyGlnAlaIleGluTyrSerGlyLeuThrGlnThrGlnLeuArg	508
Db	2759	GGTATGTGTTTATTAGACACTGCCACATTTCTACACGCGCTAAACGTAGAGATGAACCA	2818
OY	509	AlaAlaThrGlnLeuAlaIleTyrTyrPheThrAsp-----SerAlaGluLeu	524
Db	2819	GCAATTCGGGAAGTTTCTTTGAACATAATTGATATGTACAGCATACAGCAACGCAATTT	2878
OY	525	AspLysAspLysLeuLysAspTyrHisGlyPheGlyAspMet--AsnAspSerThrLeu	543
Db	2879	ACAACAGATGAAAAAGGTCAATCTCCCTTGATGCCACATGCACAGCATGATTAATCTTTG	2938
OY	544	AlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProGlnLeuThr---	562
Db	2939	CGAGTAACGAATGTACCGCAGGAATATTCGCTGGTGAAGAG-----TATTTGACAGGA	2992
OY	563	-----AspLeuAspPhePheIlePro-----	569
Db	2993	AAAGCATTAAGCTGTGTCAAAGAGACACAAACAACTAAATTCATTACAGAAAAACATT	3052
OY	570	AsnAsnAsnLysTyrGln-----SerLeuIleGlyThrGlnThrHisPro	584
Db	3053	GATCACACACTGTTTACAGATCAAGATTCACAGATTATGTGTCGGCGGATTCATGGAACCA	3112
OY	585	GluAspLeuValAspIleIleArgMetLysAspLysLys---GluValIleProValThr	603
Db	3113	GAAAGG-----AAGTTTGTTCAGCAACACATTAACAGGTCAGACGTCCTTCGAA	3166
OY	604	HisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaGlyAspArgThrLysAspPhe	623
Db	3167	-----AAATACAGTGTTCAGCTCAAGTTGATTAACANCAAAAGCAGGCTT	3211
OY	624	HisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThr	643

Db	3212	TATCCAAATTATTTCACGAGCAGCAAGGTAAGAACAACAGCCGTTATGACCGTCAAAACC	3271
Oy	644	AsplystrhranleugluPhelLysAspGlyLysalathrLleasleuLysHsiGlyglu	663
Db	3272	GACCAACCTAAGTAGGAGGTCAAAGAT-----ACAACGATTATGTT-----GGTGT	3319
Oy	664	SerLeuThrLeuGingluLeuProgluGlyTrSerGlyLeuValLysgluThrAspSer	683
Db	3320	TCGtGGGAAA-----CCAGAA-----GATAATTTCGTTTCACGCACAGACAA	3361
Oy	684	GluglyTyrylVal-----LysValAsnSerGingluValAlaAsnAlaThrVal	700
Db	3362	ACAGGTCACAGCGTCCGTTGAAAAAATGATGTCACGGGACACAGTGAAT-----GTT	3415
Oy	701	SerLysThrGlyLleThrSerAspGluThrLeuAlaPheGluu	715
Db	3416	GATTAATAATGAGC-----GATTATGCAATTGCTATATAAAT	3451

RESULT 12

ID	ABN98186	standard; DNA; 4249 BP.
...		

AC ABN98186;

DT 05-AUG-2002 (first entry)

DE E faecalis EF104 gene fragment.

KW Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic;

KW gene; ds.

OS Enterococcus faecalis.

PN US2002045737-A1.

PD 18-APR-2002.

PF 04-MAY-1998; 98US-0071035.

PR 04-MAY-1998; 98US-0071035.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Choi GH, Bailey C, Hromockyj A, Kunsch CA;

DR WPI; 2002-425450/45.

XX 5
7
8227
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
84

PT vaccines for preventing, treating or attenuating an infection caused by

[illegible]

PS Claim 1; Page 183-185; 255pp; English

CC The present invention provides the protein and coding sequences of a

CC used as vaccines for preventing or attenuating an infection caused by a

CC The polynucleotide is also useful for preventing or treating E. faecalis

XX

SQ Sequence 4249 BP; 1579 A; 779 C; 815 G; 1069 T; 7 other;

Alignment Scores:

Score: 161.00

Best Local Similarity:	21.06%	Mismatches:	326
------------------------	--------	-------------	-----

DB:	24	Gaps:	40
-----	----	-------	----

US-09-494-297-2 (1-757) x ABN98186 (1-4249)

OY	39	AlaLeuValIThrSerMetValGlyAlaLysThrValPheGlyLeuValGluSer-----	56
Db	1451	TCATTAAATACACCGTGAATGGCCCTAATTAAGCAATTCATTAAGTATCCGATCAATAT	1510
OY	57	---SerThrProAsnAlaIleAsnProAspSerSerGluTyrArgTrpTyrGlyTyr	75
Db	1511	ATTGACCAAAATTAAGTGTGTTGAATTCCT---TTGAATGCTGAAGACTCTGGGGTAAATAT	1567
OY	76	GluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg	95
Db	1568	GATCA-----AATGGCGCTATTCATCAAGA	1594
OY	96	-----ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuLysLys	113
Db	1595	ACAACCTGCTCAGTTATGGGAAGCAAGAAACCGGATTCAAAATTTAGAAATTTAAACTA	1654
OY	114	AlaPhePro-----LeuGlySerAspSerSerValLysTyrTrpTyrLysLys	129
Db	1655	AAGATCCTCATATTATCTTTCATTTCAGACTTCACAAAGAAATTTATTTTATTATACAG---	1711
OY	130	HisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArgIleThrGlyAsp	149
Db	1712	-----TTAGGAAGC-----GATTATACAGTAAGCCCAACCTCAGATGGTTCA	1753
OY	150	GluLeuAsnGlnLysLeuArgAlaValMetTyrAsnGlnLysProGlnAsnAlaAsnGly	169
Db	1754	GTTTATTAAAGTTACTACGCCCAATACCAAGCAAAATCCAAATTCCAATTTGTTTAAATAT	1813
OY	170	IleMetGluGlyLeuGlu-----ProLeuAsnAlaIleArgValThrGln	184
Db	1814	GTGCCAGACTGTTTGGCCAAAGATAAAGATATCCAGTCGATACGATACCGATACATATG	1873
OY	185	GluAlaValIlePtyTyrTyrSerAspAsnAlaProIleSerAsnProAspGluSerPheLys	204
Db	1874	AGTCTCTGAAGCTTTAACTCCAGTGTATPCGACAGTAACTACT-----AATAGTAAAG	1924
OY	205	ArgGluSerGlnSerAsnLeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLeu	224
Db	1925	CGTGCTTGTGAACACACACTTCAAGATGTATTAATCAATCTCTTGTAATGCACGAAT	1984
OY	225	LysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAspPhe	244
Db	1985	GATTCCTTTGACTCAGTACGACGTCGCGTCCAAAAATTTCCAGCTGGCGCC-----GATGTT	2038
OY	245	GlnLeuSerIlePheGluSerGluAspLysLysAspTyrAsnLysGlyTyrGlnAsn	264
Db	2039	CTTTTGTGACATTTATGATGTTTCAAC-----GATCAGGTAGATTCATTAATATCCACA	2092
OY	265	LeuLeuSerGlyGlyLeuValProThrLysProProThrProGlyAspProPheMetPro	284
Db	2093	TACTGGAGACCCGGTCAATACTTGTGATTAACCATGACGCCAACACGCCCTGGATATCCA	2152
OY	285	-----ProAsnGlnProGlnThrThrSerValLeuIle-----Arg	296
Db	2153	ACGATTACTTTTGAAGAAATACCAATATGATTACAGTTTGGATTTTGGAAAAACCAACAAA	2212
OY	297	LysTyrAlaIleGlyAspTyrSerLysLeu-----LeuGluGlyAlaThrLeuGln	313
Db	2213	CGTTACATTAATTT---GAGTATTAATAAACGCCCAATGCGTGGATCGACGTGCCAATCTTTAT	2268
OY	314	LeuThrGly-----AspAsnValAsnSerPheGlnAlaArgValPhe	327
Db	2270	ATAACAGGACAGCAAGAAAGAACCAACATTCATATATATGAAGGCTGCTCGTTCCT	2329
OY	328	SerSerAsnApIleGlyGluArgIleGluLeuSerAspGlyThrTyrThrLeuThrGlu	347
Db	2330	GTTCAAAAT-----GAAAGCCTTACACTTTTGAAGTCGACAA-----	2365
OY	348	LeuAsnSerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAlaGly	367
Db	2366	-----CAAGCGGCGGATCCACATTTAAAAAATGTATACAAAACG	2404

OY		368	LysValIlyTrThr-----lleileasprcllys-----glnlllelgluasproasnlysclu	384
Db		2405	ACAGcTAAACAACCAAAAAATATTGATATAAACACATCGCTGGTGAACAAAACGATTA	2466
OY		385	IleValGIuProTySerValGIuLalylarAsnAprhelegluLupheservalLeuthr	404
Db		2455	TTAACAACCAAAAGGCACAAACCANTGCTCAATTCGATTGTTCATTACCSTG-----	2518
OY		405	ThrgInasnTyAlaLalysPheTyTrAlalaysAsnGlySerSerglinVal	424
Db		2519	-----AAAGCGTGCCAGAAGATCCT	2539
OY		425	TyrCySPheasnAlaAspleuLySerProProaspSergluAsprGlyLystrMet	444
Db		2540	TAT-----TCATTAGAGAAAGCATACAAACGGTGGCGCAATCATT	2578
OY		445	ThrProaspPheThrThrglyLu-----ValLystrYthrHisllealelaryg	461
Db		2579	TTTAAAGCTATACATTGACAGAAACATTACGATTGAATACATACAGTCTCTGCCAAC	2633
OY		462	AspleuPheLystrYThrValLySproArgspThrAspProaspThrPheleuShis	481
Db		2639	GCTGGCCAATCTATACAGAAACAAACAAATCGACTGAAACATGTAACCAAGATGCTGCT	2699
OY		482	IleLySlyValIle-----gluLys	488
Db		2699	AGCAAAGAAAAAGTCACCACCTGGCCAAATCACATTGAATTCCTCAAGAGTGATCCGGA	2758
OY		489	GlyTyTrAgsluLysGLyGlnAlallegluyrserGlyleuThrglyThgleuLarg	508
Db		2759	GGTATTGTTTATTAGCAACTGCCACATTCTACACGCATTAACGTAGAGATGAACCA	2818
OY		509	AlalaThrGlnleuAlalleTyTyrrPheThrAsp-----SerAlagluLeu	524
Db		2819	GCAATTGGCAAGGTTCTTTGAACATAATTGATATGTCACGCATACAGAACGCAATTT	2878
OY		525	AsplysaspyLysleuLysAspyTrHisqlPheglyasPmet--AsnapSerThreu	543
Db		2879	ACAAACGATGAAAAAGTCCAATCTCTTGATGCCATCAGACAGGTGAATATCTTTG	2938
OY		544	AlaValAlalySilleleuValGIuLyrrAlaglnasPserAsnProrogLleuthr---	562
Db		2939	CGAGTAACGAATGTACCGAGAAATATTCGCGATGAGAG-----TATTTCACAGCA	2999
OY		563	-----AspleuasPhePhelePro-----	569
Db		2993	AAAGCATTAAAGCTGCTCAAGAGACACAAACACTAAAAATTCCTTACCAAAAACATTT	3053
OY		570	AsnAsnAsnLystrYTrGln-----SerleuIlleglyThrglnTrhisPro	584
Db		3053	GATCCACAGCTCGTTTCAAGTCAAGGATTCACAGATTTATGTGCGGATTCATGGAAACCA	3112
OY		585	GluaSprleuValAspIlelelarymetsluAsprlySlys---GluValIIeProvalThr	603
Db		3113	GAAAGG-----AACTTTGTTTCACCAACAGATTAACAGGTCAAGACGTTGCTTGC	3166
OY		604	HisaSnleuthrLeuArgTrhyValThrglyleuAlaglAspArgThryLysAprhe	623
Db		3167	-----AAATATCATCTTTTCAGSTCAAGTTGATTAACANCAAAAGCAGCCTT	3211
OY		624	HisPheglullegluleuLysAsnAsnlysglnGluLeuLeuSerglinThryValLystr	643
Db		3212	TATCAAAATTATACAGTAGACAGAAAGTAAAGACAAAGACGCTATGTACGCTAAACCC	3272
OY		644	AspyLsthrAnleuGluPheLysAsprGlyLysAlatrIleasnleuLynIslelyLu	663
Db		3272	GACCAATGTAAAGTAGAGCTCAAAAGAT-----ACAAACGATTATGTT-----GCTAT	3319
OY		664	SerleuthrHleuglnGlyleuProglinLytyrSeryrleuVallyGlnThrasPser	683
Db		3320	TCGTGGAAA-----CCAGAA-----GATTAATTTCGTTTCAGCGACAGACAA	3361
OY		684	GluclyTyrrLySlyVal-----LysValasnSerglnGluValAlaasnAlathrVal	700

DB 3362 ACAGTCACGACGCGCTTGAAGAAATTCATGCTTCAGCAACAGTGAAT-----GTT 3415
QY 701 SerLysThrGlyIleThrSerAspGluThrLeuAlaIleGluAsn 715
DB 3416 GATTAATAATAGGC-----GATTATGAATAATGCTATATAAAT 3451
RESULT 13
ABX61756
ID ABX61756 standard; DNA: 4249 BP.
AC ABX61756;
XX
XX 26-FEB-2003 (first entry)
DT
XX Enterococcus faecalis EF040 polynucleotide #202.
DE
XX EF040: gene; ds; immunostimulant; antibacterial; gene mapping.
XX
OS Enterococcus faecalis.
XX
XX US6448043-B1.
PN
XX 10-SEP-2002.
PD
XX 04-MAY-1998; 980S-0071035.
PF
XX 06-MAY-1997; 970S-044031P.
PR 16-MAY-1997; 970S-046655P.
PR 14-NOV-1997; 970S-066099P.
PR 14-NOV-1997; 970S-066099P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX ChOI GH, Bailey C, Hromockyj A, Kunsch CA;
XX WPI; 2003-089120/08.
DR P-PSDB; ABU13709.
DR
XX
XX New EF040 polypeptides and polynucleotides from Enterococcus faecalis.
PT useful for generating an immune response against E. faecalis and other
PT Enterococcus species, and as vaccines against other bacterial genera -
PT
XX
XX Example 1; Column 205-208; 146pp; English.
XX
XX The invention relates to polynucleotide fragments of a gene from
CC Enterococcus faecalis, EF040, and the polypeptides encoded by them. The
CC polypeptides are useful in detecting E. faecalis, as epitope tags, as
CC molecular weight markers on SDS-PAGE gels or for molecular sieve gel
CC filtration columns, in generating antibodies that specifically bind to
CC the E. faecalis polypeptides, in generating an immune response against E.
CC faecalis and other Enterococcus species and as vaccines against other
CC bacterial genera. The polynucleotides are useful as probes for gene
CC mapping and for identifying E. faecalis in biological samples. Sequences
CC ABX61555-ABX61802 represent EF040 polynucleotides of the invention.
CC Note: the sequence data for this patent can also be obtained from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
XX
SQ Sequence 4249 BP; 1579 A; 779 C; 815 G; 1069 T; 7 other:

Alignment Scores:
Pred. No.: 0.00911 Length: 4249
Score: 161.00 Matches: 159
Percent Similarity: 34.83% Conservative: 104
Best Local Similarity: 21.06% Mismatches: 326
Query Match: 4.08% Indels: 166
DB: 25 Gaps: 40

US-09-494-297-2 (1-757) x ABX61756 (1-4249)
QY 39 AlaleuValThrSerMetValGlyAlaLysThrValPheGlyLeuValGluSer----- 56
DB 1451 TCATTAACTACACCTGTAATTTGCTCTAATTAAGCCATTCAATTAGTATCCGATCAATAT 1510

QY 57 ---SerThrProAsnAlaIleAsnProAspSerSerGluTyrArgTyrGlyTyr 75
DB 1511 ATTAACCAATTAAGTGTGTTAATCCT---TTGATGCTGAACCTGTTGGGTAATAT 1567
QY 76 GluSerThrValArgGlyHisProTyrTyrGlnPheArgValAlaHisAspLeuArg 95
DB 1568 GATCAA-----AATGGTGCCTATTCATCAAGA 1594
QY 96 ---ValAsnLeuGlySerArgSerTyrGlnValTyrCysPheAsnLeuValLys 113
DB 1595 ACAACTGCTCAGTTATGGAGCAAGCAAAACCGATTCATAATTTAGAAATTAAGTA 1654
QY 114 AlaPhePro-----LeuGlySerAspSerValLysLysTyrLysLys 129
DB 1655 AAGCATCTTAATTAATCTTCTTATTCAGAGCTACAAAGAAATTTATTTTACAG--- 1711
QY 130 HisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArgIleThrGlyAsp 149
DB 1712 ---TTAGAGACG-----GATTATACAGTACGCCACGTCAGATGTTCA 1753
QY 150 GluLeuAsnGlnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnGly 169
DB 1754 GTTATTAACTTCACTACGCCAATTAACCAAGAAATCCAAATTCGTTTAAATAT 1813
QY 170 IleMetGlnGlyLeuGlu-----ProLeuAsnAlaIleArgValThrGln 184
DB 1814 GTGCCAGATAGTGTGGCAAAAGATAAGATATCCAGTCCATACGATACCGATACATG 1873
QY 185 GluAlaValIlePyrTyrSerAspAsnAlaProIleSerAsnProAspGluSerPheLys 204
DB 1874 AGTGCCTGAAGTTAATCCAGTTGATACGACAGTACTACT-----AATAGTAAG 1924
QY 205 ArgGluSerLysSerAsnLeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLeu 224
DB 1925 CGTGGTTCTGAAGCAACACTTCAAGTAGTAATAATCAATTCCTTGCAATGACGAAAT 1984
QY 225 LysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAspPhe 244
DB 1985 GATTCCTTTGACACACACGACGTCGCCGCAAAATTCACCTGCGGCC-----GATGTT 2038
QY 245 GluLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsnLysGlyTyrGlnAsn 264
DB 2039 CTTTGTGACATTTATGATGTTTCAAAC-----GATCAGGTATGATCAATTTTCCACAA 2092
QY 265 LeuLeuSerLysGlyLeuValProThrLysProProThrProGlyAspProPheMetPro 284
DB 2093 TACTGGACCGCGGTCAATCTTGTGATTAACCAATGACGCCAAACACCCCTGATATCCA 2152
QY 285 ---ProAsnGlnProGlnThrThrSerValLeuIle-----Arg 296
DB 2153 ACGATTACTTTTACGCAAAATACCAATAGTTACAGCTTATTTGGAAACCAACAA 2212
QY 297 LysTyrAlaIleGlyAspTyrSerLysLeu-----LeuGlnGlyAlaThrLeuGln 313
DB 2213 CGTTACATTTATTT---GAGTATTAATAACGCCAATGCTGATGACGCGCCACCTTTTAT 2269
QY 314 LeuThrGly-----AspAsnValAsnSerPheGlnAlaArgValPhe 327
DB 2270 ATTAACGGGACACGCAAGAACCAATGCAATGATTAATGTAAGCGCTGCTTGCGTTTCT 2329
QY 328 SerSerAsnAspIleGlyAlaArgIleGluLeuSerAspGlyThrThrLeuThrGlu 347
DB 2330 GTTCAAAAT-----GAAGGCTTAAGACATTTTGATGTCACAA----- 2365
QY 348 LeuAsnSerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAlaGly 367
DB 2366 ---CAAGCGCGCAATTCACCAATTAATAAATAAATAAATAAAGCAAGCAAG 2404
QY 368 LysValTyrThr---IleIleAspGlyLys-----GlnIleGluAsnProAsnLysGlu 384
DB 2405 ACAGTAAACAACAAATAATTTGATTAATAAACACATCGTGTGAATAATCCACGATGAA 2464

```

OY 385 ILeValGluProTyrSerValGluValTyrAsnAspPheGluGluPheSerValLeuThr 404
DB 2465 TTAACACCAAAAGGCACAAACCAATCGTCAATGCTTGAATTCATTATACCGTGTG----- 2518
OY 405 ThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAsnLysSerGlnValVal 424
DB 2519 -----AAAGGGTCCGACAAATGCT 2539
OY 425 TyrCysPheAsnAlaAspLeuLysSerProAspSerGluAspGlyLysThrMet 444
DB 2540 TAT-----TCATTAGAGAACTACAAAGCTGCCAAAGTCAT 2578
OY 445 ThrProAspPheThrThrGlyGlu-----ValTyrThrHisIleLeuArg 461
DB 2579 TTTAAAGACTTATCTATACAGAAACATTCAGATGATTAATACGCTCTCCCAAC 2638
OY 462 AspLeuPheLysTyrThrValLysProArgAspThrAspProAspThrPheLeuLysHis 481
DB 2639 GCTGCCAAATCTATACAGAAACCAACATCGACTGAAACATGACAGATGCTGCT 2698
OY 482 IleLysLysValIle-----GluLys 488
DB 2699 AGCAAGAAAAAGTCACACATGCCCAATCACATTGAATTCTCAGAAAGTCATGCCGAA 2758
OY 489 GlyTyrArgGluLysGlyGlnAlaIleGluTyrSerGlyLeuThrGluThrGlnLeuArg 508
DB 2759 GGTATGTTTATTATAGCAACCTGCCACATTCACGCAATACGTTAGAGATGAAACCA 2818
OY 509 AlaAlaThrGlnLeuAlaIleTyrTyrPheThrAsp-----SerIleGluLeu 524
DB 2819 GCAATTTGCCAAGGTTCTTTTGAACATAATGATATGTCACGCAATACAGCAACGAAATT 2878
OY 525 AspLysAspLysLeuLysAspTyrHisIleGlyAspMet---AsnAspSerThrLeu 543
DB 2879 ACAACAGATGAAAAAGTCAATACTCTTTCATGCCATGACACAGCTGATATATCTTGG 2938
OY 544 AlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAspProGlnLeuThr--- 562
DB 2939 CGAGTAACGAATGACCGCAGGAATATTCCTGGATGAAAGAG-----TATTGACAGGA 2992
OY 563 -----AspLeuAspPhePheIlePro----- 569
DB 2993 AAAGCCATTAAAGCTGTCAAAGAGACAAACCACTAAATTCATTAAACGAAACAAATT 3052
OY 570 AsnAsnAsnLysTyrGln-----SerLeuIleGlyThrGlnTrpHisPro 584
DB 3053 GATCCAGCTCGTTTACAGTCAAGTCAACAGATTCAACGATTTATGCGCGATTCATCAAGAACCA 3112
OY 585 GluAspLeuValAspIleIleArgMetGluAspLysLys---GluValIleProValThr 603
DB 3113 GAAAG-----AACTTTGTTTCAGACACAGATTAACAGATTCAGACGTTCCCTTCGAA 3166
OY 604 HisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaGlyAspArgThrLysAspHe 623
DB 3167 -----AAATCACGTTTCAGGTCAGTCAAGTTGATTAACANCAACAGAGCGTT 3211
OY 624 HisPheGluIleGluLeuLysAsnLysGlnGluLeuLeuSerGlnThrValLysThr 643
DB 3212 TATCCAAATTAATTACAGTACGAGAAAGTAAAGAAAGAGCCATATGTGACCGCTCAACCC 3271
OY 644 AspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyGlu 663
DB 3272 GACCAATCATTAAGTACAGTCAAAAGAT-----ACAAACGATTATGTT-----GCTGAT 3319
OY 664 SerLeuThrLeuGlnGlyLeuProGlnGlyTyrSerTyrLeuValLysGluThrAspSer 683
DB 3320 TCGTGGAAA-----CCAGAA-----GATTAATTTCGTTTCACGCGACACAA 3361
OY 684 GluGlyTyrLysVal-----LysValAsnSerGlnGluValAlaAsnAlaThrVal 700
DB 3362 ACAGGTCAAGACGTCCTTCGTTGAAGAAATTTGATGTTTCAGGCAACAGTGAAT-----GTT 3415
OY 701 SerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAsn 715

```

```

DB 3416 GATAAAATRAGC-----GATTATGAATTTGCTATATAAAT 3451
RESULT 14
ID AAX20200
AXX20200 standard; DNA; 4359 BP.
XX
AC AAX20200;
XX
DT 20-APR-1999 (first entry)
XX
DE Enterococcus faecalis gene EF104.
XX
KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX detection; attenuation; antigenic; ss.
XX
OS Enterococcus faecalis.
XX
MO W09850554-A2.
XX
PD 12-NOV-1998.
XX
PF 04-MAY-1998; 98WO-US08959.
XX
PR 14-NOV-1997; 97US-0066009.
XX 06-MAY-1997; 97US-0044031.
XX 16-MAY-1997; 97US-0046653.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bailey C, Chol GH, Hromockyj A, Kunsch CA;
XX
DR WPI: 1999-070095/06.
XX
DR P-PSDB: AAY0210.
XX
PT New isolated Enterococcus faecalis polynucleotides - used to develop
PT products for the detection of Enterococcus and for use in vaccines
PT for prevention or attenuation of Enterococcus infection
XX
PS Claim 1; Page 200-201; 301pp; English.
XX
CC The present sequence represents a gene isolated from
CC Enterococcus faecalis. The present invention describes genes, proteins
CC and antigenic polypeptides isolated from E. faecalis. The proteins can
CC be used in vaccines for preventing or attenuating an infection caused
CC by a member of the Enterococcus genus in an animal. They can also be
CC used for detecting Enterococcus antibodies in a sample. The nucleotide
CC sequences can be used for detecting Enterococcus nucleic acids.
CC products from the present invention can also be used for screening
CC compounds to identify agonists and antagonists of E. faecalis protein
CC activity.
XX
SQ Sequence 4359 BP; 1614 A; 794 C; 838 G; 1106 T; 7 other;
XX
Alignment Scores:
Pred. No.: 0.0094 Length: 4359
Score: 161.00 Matches: 159
Percent Similarity: 34.83% Conservative: 104
Best Local Similarity: 21.06% Mismatches: 326
Query Match: 4.08% Indels: 166
DB: 20 Gaps: 40
US-09-494-297-2 (1-757) x AAX20200 (1-4359)
OY 39 AlaLeuValThrSerMetValGluValLysThrValPheGlyLeuValGluSer----- 56
DB 1546 TCATTATAGTACACCGTGAATTTGCTCTTAATAAGCCATTCAATTATGATCCGATCAATAT 1605
OY 57 --SerThrProAsnAlaIleAsnProAspSerSerSerGluTyrArgTyrTrpGlyTyr 75
DB 1606 ATTGAACCAATATAGTGTGTTAATCCT--TTGAATGCGTGAAGACGCTGGGATTAATAT 1662
OY 76 GluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95

```

```

Db      1663 GATCAA-----ATGTCCTCATTCATCAAGA 1689
Qy      96 -----ValasnleugllyserArgSerTyrGlnValTyrGlyPheasnLeuLys 113
Db      1690 ACAACTGTCTCATGTTATGGAGCAAGAACCGATTCATAAATTTAGAAATTAAGTA 1749
Qy      114 AlaPhePro-----LeuGlySerAspSerSerValLysLysTyrLysLys 129
Db      1750 AAGATCCTTAATTATCTTTCATTACGAGCTACAAAACAATTTTATTATACAG--- 1806
Qy      130 HisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArgIleThrGlyAsp 149
Db      1807 -----TTAGAACG-----GATTATACAGTACGCCACGTCAGATGGTTCA 1848
Qy      150 GluLeuasnGlnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnGly 169
Db      1849 GTTATTAGTTCACCTACGCCAATACCAACGAATCCAAATTCCAATGGTTTAAATTAT 1908
Qy      170 IleMetGluLysLeuGlu-----ProLeuasnAlaIleArgValThrGln 184
Db      1909 GTGCACAGATAGTTTGGCAAAAAGATAAAGATATCCAGTCGATACGATACGATACAAATG 1968
Qy      185 GluAlaValTrrPyrTyrSerAspAsnAlaProIleSerAsnProAspGluSerPheLys 204
Db      1969 AGTCGCTGAAGTTTAACTCCAGTTGATACGACAGTAACTACT-----AATAGTAG 2019
Qy      205 ArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLeu 224
Db      2020 CGTGTTCTGACAGCAACACTTCAAAAGTAGTAAATAATTCCTTGTCATGCGCAAAAT 2079
Qy      225 LysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAspPhe 244
Db      2080 GATTCTTTTACTACACTACAGTCGCCGACAAAATTCACAGTCGCGCC-----GATGTT 2133
Qy      245 GluLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsnLysGlyTyrGlnAsn 264
Db      2134 CTTTTCGACATTTATGATGTTTCAAAAC-----GATCAGTAGATTCATTAATTTATCCAA 2187
Qy      265 LeuLeuSerGlyLysLeuValProThrLysProProThrProGlyAspProPhePro 284
Db      2188 TACTGGAGCGCGCTCAACTCTTGTATAAACCAATGACGCCAAACACCCCTGGATATCCA 2247
Qy      285 -----ProasnGlnProGlnThrThrSerValLeuIle-----Arg 296
Db      2248 ACGATTACTTTTACGAAATAATACCAATAGTTACAGTTGATTTTGAAAAACCAACAAA 2307
Qy      297 LysTyrAlaIleGlyAspTyrSerLysLeu-----LeuGluGlyAlaThrLeuGln 313
Db      2308 CGTTACATTTATTT--GAGTATTAATAAACGCCAATGCTGATCGACGTGCCAATCTTTAT 2364
Qy      314 LeuThrGly-----AspAsnValAsnSerPheGlnAlaArgValPhe 327
Db      2365 ATTAACAGGACAGCAAGAAAGAACCAATCAATATATATGAGGCTCTGCTGGGTTCT 2424
Qy      328 SerSerAsnAspIleGlyLysArgIleGluLeuSerAspGlyTyrThrLeuThrGlu 347
Db      2425 GTTCAAAAT-----GAAGCGTTAGACATTTTGATGCAACA----- 2460
Qy      348 LeuAsnSerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGluAlaGly 367
Db      2461 -----CAAGCGGCGAATCCAAATTAATAAATGTAACAAAACG 2499
Qy      368 LysValTyrThr-----IleIleAspGlyLys-----GlnIleGluAsnProAsnLysGlu 384
Db      2500 ACAGTAAACAACAAAATATATGATATTAATAAACACATCTGTGCAAAAATCCACAGATTGAA 2559
Qy      385 IleValGluProTyrSerValGluAlaTyrAsnAspPheGluGluPheSerValLeuThr 404
Db      2560 TTAAACACCAAAAGGACAAACCAATGCTCAATCGATTGAATCTATACCGG----- 2613
Qy      405 ThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAsnLysAsnGlySerSerGlnValVal 424

Db      2614 -----AAAGCGTCGCAAGATGCT 2634
Qy      425 TyrCysPheAsnAlaAspLeuLysSerProProAspSerGluAspGlyLysThrMet 444
Db      2635 TAT-----TCATTAGACAGACTCAACAAAGGTGGGAAGTCAATT 2673
Qy      445 ThrProAspPheThrThrGlyGlu-----ValLysTyrThrHisIleAlaGlyArg 461
Db      2674 TTTAAGACATATACATTGACAGAAAACATTACGATGTAATACATACAGCTCTGCAAAAC 2733
Qy      462 AspleuPheLysTyrThrValLysProArgAspThrAspProAspThrPheLeuLysHis 481
Db      2734 GCTGGCCAAATCTATACAGAAACAACAAATGACTGTGAACATTTGAACAGAGTCTGCT 2793
Qy      482 IleLysValIle-----GluLys 488
Db      2794 AGCAAGAAAAGTCACCACTCGCCCAATACATTGAATTTCCAGAAAGTGATCGGAA 2853
Qy      489 GlyTyrArgLysGluAlaIleGluTyrSerGlyLeuThrGluThrGlnLeuArg 508
Db      2854 CGTATGTTTATTATAGCAACTCGCCACATCTACACGCAATACGTAGAGATGAACCA 2913
Qy      509 AlaAlaThrGlnLeuAlaIleTyrTyrPheThrAsp-----SerAlaGluLeu 524
Db      2914 GCATTTCGGAAGTTCTTTTGAATGATGATATGTCACGCATACGACGACCGAATTT 2973
Qy      525 AspLysAspLysLeuLysAspTyrHisGlyPheGlyAspMet---AsnAspSerThrLeu 543
Db      2974 ACAACATGTAAGAAAGTCATCTCTCTTGATGCCATGACAGAGGTGTTATCTTTC 3033
Qy      544 AlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProProGlnLeuThr--- 562
Db      3034 CGAGTACGATGATACGCGACGAGATATTCGCTGATGAAAG-----TATTGACAGCA 3087
Qy      563 -----AspleuAspPheThrLeuPro----- 569
Db      3088 AAAGCCATTAGCTGTCAAAGACAGACCAACCACTAAATAATTCATTAAACGAAACAAATT 3147
Qy      570 AsnAsnAsnLysTyrGln-----SerLeuIleGlyThrGlnTrrPheLysPro 584
Db      3148 GATTCACGTCGTTTACAGCAAGTCAAAAGATTCACATTAATTCGCGCTTATGGAACCA 3207
Qy      585 GluAspLeuValAspIleIleArgMetGluAspLysLys---GluValIleProValThr 603
Db      3208 GAAGAG-----AACTTTGTTTACGACAACAGATTAATAAACAGTCAACAGCTTCCCTTCA 3261
Qy      604 HisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaGlyAspArgThrLysAspPhe 623
Db      3262 -----AAATTCAGCTGTTTCAGGTCAAGTGTATACATCAACAAAGACGCCGTT 3306
Qy      624 HisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThr 643
Db      3307 TATCCATTTTATTACACTGACGAAGTAAGAAAGAAACACCCATATGACGCGCAAAACC 3366
Qy      644 AspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyGlu 663
Db      3367 GACCAATCTAAGTTAGAGGTCAAAGAT-----ACAACGATTTATGTT-----GGGAT 3414
Qy      664 SerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGluThrAspSer 683
Db      3415 TCGTGCAAA-----CCAGAA-----GATTAATTCGTTTCAGCGACGACAAA 3456
Qy      684 GluGlyTyrLysVal-----LysValAsnSerGlnGluValAlaAsnAlaThrVal 700
Db      3457 ACAGGTCAACACGTCGTTTGAATAAATGATGTTTCAGGACGACAGTAAAT-----GTT 3510
Qy      701 SerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAsn 715
Db      3511 GATTAATAATAGC-----GATTATGAATAATGTCTATAAAT 3546

RESULT 15
ABN98185
ID      ABN98185 standard; DNA: 4359 BP.

```



```

XX ABN98185;
AC
XX
XX 05-AUG-2002 (first entry)
DT
XX
XX E faecalis EPI04 gene.
DE
XX
XX Enterococcus; vaccine; gastrointestinal disease; diagnosis; antibiotic;
KM gene; ds.
XX
XX Enterococcus faecalis.
OS
XX US2002045737-A1.
XX
XX 18-APR-2002.
PD
XX
XX 04-MAY-1998; 98US-0071035.
PF
XX
XX 04-MAY-1998; 98US-0071035.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Choi GH, Bailey C, Hromockyj A, Kunsch CA;
PI
XX
XX WPI; 2002-425450/45.
DR
XX
XX P-PSDB: ABP43429.
XX
XX
XX New genes and polypeptides from Enterococcus faecalis, useful as
PT vaccines for preventing, treating or attenuating an infection caused by
PR a member of the Enterococcus genus in an animal, particularly E.
PT faecalis
XX
XX
XX Claim 1: Page 180-182; 255pp; English.
PS
XX
XX
XX The present invention provides the protein and coding sequences of a
CC number of polypeptides from Enterococcus faecalis. The proteins can be
CC used as vaccines for preventing or attenuating an infection caused by a
CC member of the Enterococcus genus in an animal, particularly E. faecalis.
CC The polynucleotide is also useful for preventing or treating E. faecalis
CC infection. The present sequence is a coding sequence of the invention.
CC
XX
XX
XX Sequence 4359 BP; 1614 A; 794 C; 838 G; 1106 T; 7 other;
SQ

```

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:	Query Match:	Gaps:
0.0094	4359	161.00	159	104	21.06%	326	166	4.08%	40

US-09-494-297-2 (1-757) x ABN98185 (1-4359)

```

OY 39 AlaleuValThrSerMetValG1ValAlaYsthrValPheGlyLeuValG1user----- 56
   ||||| ||| ||||| ||| ||||| |||||
DB 1546 TCATTAACTACACCTGTAATGGTCTGCTAATAAGCATTCATTGTAATCCGATCAATAT 1605
   ||||| ||||| ||||| ||||| ||||| |||||
OY 57 ---SerThrProAsnAlaIleAsnProAspSerSerGluTyrArgTrpTyrG1Ytyr 75
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1606 ATTGAAACCAATTAGTGTGTTAATCCT---TTGAATGCTGAAACGTCCTGGGTAATAT 1662
   ||||| ||||| ||||| ||||| ||||| |||||
OY 76 GluserTyrValArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1663 GATCAA-----AATGCTGCTATTCATCAAGA 1689
   ||||| ||||| ||||| ||||| ||||| |||||
OY 96 -----ValAsnLeuGlySerArgSerTyrGlnValTyrCysPheAsnLeuLysLys 113
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1690 ACAACTGCTCAGTTATGGGAAGCAAGAAACCGATTCAAAATTTGAATATTAAGTA 1749
   ||||| ||||| ||||| ||||| ||||| |||||
OY 114 AlaPhePro-----LeuGlySerAspSerSerValLysLysTrpTyrLysLys 129
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1750 AAGCATCTCTAATTATCTTTCAATTACGACCTACAAAGAAATTTATTTATTACAAAG--- 1806
   ||||| ||||| ||||| ||||| ||||| |||||

```

```

OY 130 HisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArgIleThrGlyAsp 149
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1807 -----TTAGGAACG-----GATTATACAGTAACGCAACGTGACGTCTCA 1848
   ||||| ||||| ||||| ||||| ||||| |||||
OY 150 GlueAsnGlnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnGly 169
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1849 GTTATTAGTTCACATACGCAATTAACCAAGAAATCCAAATTCGAATTTGTTTAATAT 1908
   ||||| ||||| ||||| ||||| ||||| |||||
OY 170 IleMetGlyGlyLeuGlu-----ProLeuAsnAlaIleArgValThrGln 184
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1909 GTGCCAGATAGTTTCCAAAAGATAAAGTATACCGATCCGATACGATCCGATCAATG 1968
   ||||| ||||| ||||| ||||| ||||| |||||
OY 185 GluAlaValTrpTyrTyrSerAspAsnAlaProIleSerAsnProAspGluSerPheLys 204
   ||||| ||||| ||||| ||||| ||||| |||||
DB 1969 AGTGCTGAAGTTTACTTCCAGTTCATACGACGATCACTACT-----AATGCTAAG 2019
   ||||| ||||| ||||| ||||| ||||| |||||
OY 205 ArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuMetArgGlnAlaLeu 224
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2020 CGTGTCTCGAAGCAAGCACTTCGAAGTAAATCAATTCCTGTCATCCAGCAAT 2079
   ||||| ||||| ||||| ||||| ||||| |||||
OY 225 LysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGlnValProAspAspPhe 244
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2080 GATTCTTTGACTCTCACTAAGCGCTCCGTAACAAATTCACAGTCGCGCC-----GATGTT 2133
   ||||| ||||| ||||| ||||| ||||| |||||
OY 245 GlnLeuSerIlePheGluSerGluAspLysTyrAsnLysGlyTyrGlnAsn 264
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2134 CTTTGTGACATTTAAGATGTTTCAAC-----GATCAGGATGATTCATTTATTCACAA 2187
   ||||| ||||| ||||| ||||| ||||| |||||
OY 265 LeuLeuSerGlyGlyLeuValProThrLysProProThrProGlyAspProPhePro 284
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2188 TACTGGACCGCGGTCATCTTGTATTAACCAATGACGCCCAACGCCCTGATATCA 2247
   ||||| ||||| ||||| ||||| ||||| |||||
OY 285 -----ProAsnGlnProGlnThrThrSerValLeuIle-----Arg 296
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2248 ACGATTACTTTTGACGAAATATACCAATGATTCACGTTGATTTGCAAAAACCAACAA 2307
   ||||| ||||| ||||| ||||| ||||| |||||
OY 297 LysTyrAlaIleGlyAspTyrSerLysLeu-----LeuGlnGlyAlaThrLeuGln 313
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2308 CGTTACATTATTT---GAGTATTAATAAGCCCAATGCTGATCGACGTGCCCACTGTTAT 2364
   ||||| ||||| ||||| ||||| ||||| |||||
OY 314 LeuThrGly-----AspAsnValAsnSerPheGlnAlaArgValPhe 327
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2365 ATACAGGACACGCAAGAACCAACATCGAATTAATGAAGCTGCTGCTTCT 2424
   ||||| ||||| ||||| ||||| ||||| |||||
OY 328 SerSerAsnAspIleGlyGluArgIleGluLeuSerAspGlyTyrThrLeuThrGlu 347
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2425 GTTCAAAAT-----GAAGCGTTAAGACATTTGAGCTGCAACA----- 2460
   ||||| ||||| ||||| ||||| ||||| |||||
OY 348 LeuAsnSerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGlnAlaGly 367
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2461 -----CMACGCGCGAATTCACATTTAAATAATGTAAACAAAACG 2499
   ||||| ||||| ||||| ||||| ||||| |||||
OY 368 LysValTyrThr---IleIleAspGlyLys-----GlnIleGluAsnProAsnLysGlu 384
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2500 ACAGTAACACAAATAATATGATATTAATAAACACATCGTGTGAAAAAATCCAAACGATTGA 2559
   ||||| ||||| ||||| ||||| ||||| |||||
OY 385 IleValGluProTyrSerValGlnAlaTyrAsnAspPheGluGluPheSerValLeuThr 404
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2560 TTAACACCAAAAGGACACCAACCAATGCTCAAAATTCGATTGAATTCATTACCTG----- 2613
   ||||| ||||| ||||| ||||| ||||| |||||
OY 405 ThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAsnLysAsnGlySerSerGlnValVal 424
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2614 -----AAAGCGCTGCCAAGAAAGATGCT 2634
   ||||| ||||| ||||| ||||| ||||| |||||
OY 425 TyrCysPheAsnAlaAspLeuLysSerProProAspSerGluAspGlyLysThrMet 444
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2635 TAT-----TCATTAGGAAGATACCAAGCGTCCGAAGATCTT 2673
   ||||| ||||| ||||| ||||| ||||| |||||
OY 445 ThrProAspPheThrThrGlyGlu-----ValLysTyrThrHisIleAlaGlyArg 461
   ||||| ||||| ||||| ||||| ||||| |||||
DB 2674 TTTAAAGCTATACATTGACAGAAACATTAACGATTAATCAATACGCTCTGCAAC 2733
   ||||| ||||| ||||| ||||| ||||| |||||
OY 462 AspLeuPheLysTyrThrValLysProArgAspThrAspProThrPheLeuLysHis 481
   ||||| ||||| ||||| ||||| ||||| |||||

```

```
Db 2734 GCTGGCCAAATCTATACGAAACAACATGACTGTGAACATTGAACAGATCTGCT 2793
OY 482 ILeuLysValIle-----GluLys 488
Db 2794 AGCAGAGAAAAAGTCACCACTGGCCCAATCATTGAATTCTCAGAAGTGATCGGAA 2853
OY 489 GLeuTyrArgLysGlyGlnAlaIleGluTyrSerGlyLeuThrGlnLeuArg 508
Db 2854 GGTATTGTATTATTAGCACTGCCACATCTACACGATACGTAGAGATGAAACCAA 2913
OY 509 AlaAlaThrGlnLeuAlaIleTyrTyrPheThrAsp-----SerAlaGluLeu 524
Db 2914 GCAATTGCGAAGGTTCTTTGAACTAATTGATATGTCACGCAATACGCAACCAATT 2973
OY 525 AspLysAspLysLeuLysAspTyrHisGlyPheGlyAspMet---AsnAspSerThrLeu 543
Db 2974 ACAACAGATGAAAAAGTCATCTCTTGTGATGCCATCATGACAGCGATATATCTTTG 3033
OY 544 AlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProGlnLeuThr--- 562
Db 3034 CGAGTACGATGACCGCAGCAATATTCGCGTAGAGAG-----TATTGACAGGA 3087
OY 563 -----AspLeuAspPheIlePro----- 569
Db 3088 AAAGCCATTAGCTGTCAAGAGACACCACTAAATAATTCATTACGAAAAACAATT 3147
OY 570 AsnAsnAsnLysTyrGln-----SerLeuIleGlyThrGlnTyrHisPro 584
Db 3148 GATCACAGTCGTTTACAGTCAAAAGATTCAACGATTATGTGCGCGATTCATGSAACCA 3207
OY 585 GluAspLeuValAspIleIleArgMetGluAspLysLys---GluValIleProValThr 603
Db 3208 GAAGAG-----AACTTTGTTTCAGCAACAGATTAACAGGTCAAGACGTTCCCTCGAA 3261
OY 604 HisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaGlyAspArgThrLysAspPhe 623
Db 3262 -----AAATCCTGTTTCAGGTCAAGTGTATACANCAACAAAGCAGCGTT 3306
OY 624 HisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThr 643
Db 3307 TATCCATTTATTACAGTGACGAAAGTAAAGAAACAGCCTATGTGACCGTCAAAACC 3366
OY 644 AspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyGlu 663
Db 3367 GACCAATCTAAGTTAGAGGTCAAGAT-----ACAACGATTATGTT-----GCTGAT 3414
OY 664 SerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGluThrAspSer 683
Db 3415 TCGTGGAA-----CCAGAA-----GATAATTTCGTTTCAGCGACAGACAA 3456
OY 684 GluGlyTyrLysVal-----LysValAsnSerGlnGluValAlaAsnAlaThrVal 700
Db 3457 ACAGGTCACAGACGTGCCGTTTGAAAAAATGATGTTACGGACACGTGAAT-----GTT 3510
OY 701 SerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAsn 715
Db 3511 GATATAAATAGC-----GATTATGAATTTGTATATAAAT 3546
```

Search completed: August 19, 2003, 15:40:17
Job time : 580 secs